

November 1, 2006

Molly C. Dougherty, PhD, RN
Editor, *Nursing Research*
CB# 7460 School of Nursing
The University of North Carolina
Chapel Hill, NC 27599-7460

Dear Dr. Dougherty:

Attached is a revision of the article *Structural Model for Osteoporosis Preventing Behavior in Postmenopausal Women*. (Ms #2005/246). The authors were pleased with the reviewers' comments attesting to the significance and strengths of the manuscript. The specific responses to the reviewers' comments are attached. The statements of the reviewers have been carefully considered and the manuscript has been strengthened.

The attached pages are organized according to the questions and critique items of the reviewers. The reviewers' comments/critique are in bold followed by the authors' response to each item.

We hope that the manuscript is now in the form that warrants publishing in *Nursing Research*. Please contact us if there are additional questions or comments

Thank you,

Patricia J. Estok, PhD, RN, FAAN
Professor Emeriti
Kent State University
College of Nursing
Kent, OH 44242
paestok@aol.com

Reviewers' Comments and Author Responses for ms #2005/246

Manuscript Title:

Structural model for osteoporosis preventing behavior in postmenopausal women

The attached pages are organized according to the questions and critique items of the reviewers. The reviewers' comments/critique are in bold followed by the author responses to each item.

Reviewer #1 –

- **The author(s) point out the importance of osteoporosis prevention as well as the scope of the osteoporosis problem in the US. However, the purpose statement for the study does not appear until the third page of the manuscript (although we do get a hint on page two), and this should be remedied. It is not until then that the reader has a clue about what and why this study was undertaken.**

The purpose statement has been moved up in the manuscript and now appears on page 4.

- **The author(s) neglected to name the overall research design for this study. Instead, the methods section uses the analytic technique (SEM) to describe the study. It would be more appropriate to name this design as a quasi-experimental, longitudinal randomized clinical trial. SEM is the technique used for answering the specific aims of the study. I would also like to see a statement about data management and IRB compliance.**

We agree that the study is probably best described as a “longitudinal randomized clinical trial, including covariate effects.” This change has now been made.

We note that the term “quasi-experimental” is not appropriate for the current study, because it is typically reserved for studies which do not use random assignment to experimental groups (see, for example, Cook & Campbell, 1979). However, participants were randomly assigned to the two experimental conditions in the current study.

The following statement has been added to the manuscript. “Approval for the study was obtained through the Institutional Review Board, and written informed consent was obtained from all subjects and adequate provision was provided to maintain the safety and confidentiality of the subjects throughout the study.”

Reviewer #2 -

- **The manuscript would be enhanced by the addition of data on usual calcium intake and usual weight-bearing exercise in the target population.**

This information was added to the manuscript.

- **More detail should be provided on how the osteoporosis preventing behaviors were measured. What was the mean calcium intake and exercise activity of each group? Are those the values given in Table 1?**

The transformed calcium intake and exercise activity descriptive statistics for the entire sample are listed in what was previously Table 2 (and is now Table 3). However, we agree with your point that it is also very useful information for readers to know the separate mean values for each of the two groups, and also probably helpful to have this information presented in the raw (untransformed) measurement units for future reference. Based on this, we have added a new Table 1 which now reports these values. (The previous Table 1 is now Table 2).

- **Table 1 would be improved by making the title more descriptive. What is being correlated? Also, what are the units of the variables? Is 459.61 the mean calcium intake of the group in milligrams; if not, what is that number?**

The title of Table 1 is “*Mean and Standard Deviations for Daily Calcium Intake (Dietary plus Supplements) and Weekly Weight-Bearing Exercise by Experimental Group, for Times 1 through 3*”

The title of Table 2 has been changed to “*Mean, Standard Deviations and Intercorrelations Among Osteoporosis Health Belief and Knowledge Scales*”

The value of 459.61 does not refer to the calcium intake. It is the average *Calcium Self-Efficacy Scale score* for the entire sample. In reporting this result, we have preserved the original units of the survey scale (the scale has 6 items, total scores may range from 0-600). In addition, and perhaps addressing your implied larger issue, the table note for this table (now Table 2) has been expanded to indicate response options for all of the scales reported in the table.

- **On page 18, para 2, did the women have this significant bone loss while on study? Or did they have low t-scores? Low t-scores do not indicate significant loss; these women may have life-long low values compared to the reference population.**

We appreciate the reviewer’s comments and we made clear indication that these were low t-scores rather than bone loss.

- **From the standpoint of this reviewer there are too many model figures? However, these may be interesting to other readers.**

We are happy to be responsive to reviewer and editor suggestions for specific figures they feel should be cut. At this point, we felt that all figures were important to include as they either give a conceptual overview (Figure 1), or economically demonstrate both a model tested and the values of the relevant parameter estimates (Figures 2-5). However, in this revised version of the paper, we did merge old Figures 2 & 3 and Figures 4 & 7 together.

Reviewer #3-

- **There is some controversy in the literature over the effectiveness of calcium in the prevention of osteoporosis. This should be briefly discussed.**

To adequately respond to the issue would require a considerable increase in the length of the manuscript. While the authors agree that there is controversy in the literature over the effectiveness of calcium alone in the prevention of osteoporosis, it is very clear that without adequate calcium, bone maintenance and growth is jeopardized.

- **Page 3 line 6: Please provide references for the effectiveness of the prevention behaviors listed.**

The authors added one reference that encompasses the effectiveness of prevention behaviors related to osteoporosis. Citing separate research based articles for each prevention behavior would lengthen the manuscript considerably. The authors would be happy to provide a list of the numerous references available.

- **The paper does not provide a basis for the theoretical model underlying the research hypotheses. Why was knowledge and health beliefs conceptualized as a mediator of the effect of the intervention? A mediation model says that presenting people with the results of the DXA will change their knowledge and health beliefs and that this change will in turn increase OPBs. Please explain the mechanism and rationale behind this hypothesis. I wonder if moderation is more appropriate in this context. Moderation would hypothesize that providing women with the results of the DXA will lead to a greater increase in OPBs for those with high knowledge and certain health beliefs but no or a smaller increase for those with low knowledge and/or other health beliefs.**

One manner in which the intervention (DXA results) might change expressed knowledge and health beliefs is by making this information more personally relevant, thus more cognitively salient and easily (perhaps spontaneously) accessible. By providing participants with a more cohesive and self-relevant cognitive structure, information related to osteoporosis might be more organized structurally and thus scores on measures of knowledge and health beliefs might increase because the relevant information is easier to accurately retrieve.

We note, however, that our mediator model (Figure 4), does not treat either the Time 2 Health Beliefs or Knowledge constructs as mediators, although we did investigate the possibility that they could function as mediators and found no support for this idea. The empirical evidence suggests Susceptibility at T2 is potentially a mediator, although this effect was fairly weak and marginally significant. This relationship makes sense, as women who never thought of themselves as potentially having osteoporosis may have been given DXA information suggesting that they were more susceptible than they had previously believed.

- **The rationale for the design is not provided. An overview of the design is needed in the methods section. The method is described by the analysis which is somewhat worrisome. The analyses should fit the research questions and design not vise-a-versa.**

The first reviewer also raised a similar point, which we also agree with. Please see our altered descriptions of the research design and its rationale in the manuscript on pages 2 and 6.

- **Several results are reported in the methods section. Results (with the exception of reliability) should be reported in the results section.**

We have now moved substantial portions of sample-based analysis results from the methods section to the results section, and tried to organize them in a way that better presents the logic of our analyses to a reader. See page 11 for the new placement of this material.

- **Confirmatory factor analysis not principal components analysis for the OKT should be used to test if a pre-specified factor structure fits the data. Also please explain how the convergent and discriminant validity of this instrument has been supported.**

See page 14 and Figure 2 for the measurement model (CFA) results supporting our use of the OKT indicators. We did also do a confirmatory EFA at the item level to justify our sorting of OKT items into the *a priori* expected 3 subscales. These results were supportive, but are not reported in the paper due to space constraints and because they did not contradict our intended use.

The results from Confirmatory Factor Analysis of Osteoporosis Knowledge Test (OKT) are presented here. These 3 paragraphs can be added to the manuscript at the end of the section description on page 8 if the editors so desire. The authors were concerned about the length of the manuscript if this material is added.

The 24 items of the OKT were fit to a model specifying three intercorrelated latent factors. These latent factors were: (1) General osteoporosis knowledge; (2) Osteoporosis knowledge specific to exercise; and (3) Osteoporosis knowledge specific to calcium intake. The model was estimated using Mplus, version 4.1 (Muthén & Muthén, 1998-2006). A weighted least-squares estimator, yielding a mean and variance weighted chi-square statistic (WLSMV) was used, in consideration of the binary nature of the item-level responses.

The model fit adequately after allowing for six covarying uniquenesses (but no cross-loadings), with $\chi^2 = 99.826$ (df=66, p=.0045), CFI=.922, RMSEA=.048. All items loaded significantly on their intended factors. Ranges of values for the standardized loadings are as follows: (1) General, loadings of .27 to .73, with all except one loading > .40; (2) Exercise, loadings of .24 to .78, with all except one loading > .50; (3) Calcium, loadings of .41 to .76.

The General osteoporosis knowledge factor strongly related to both the Exercise factor (.88) and the Calcium factor (.85), however, the relationship of the Exercise and Calcium factors was more modest (.49). This pattern of relationships is very consistent both with past uses which have created two scale scores from the set of 24 items (Exercise scale score = General + Exercise items; Calcium scale score = General + Calcium items), and with our use in the current study in which subscale scores were used as indicators either of an Exercise Knowledge construct (2 indicators, based on sum of General items and sum of Exercise

items) or a Calcium Knowledge construct (2 indicators, based on sum of General items and sum of Calcium items).

- **A cutoff value of .09 is used for SRMR. This is a different value than what is normally used. Please provide the rationale and references for selecting .09 rather than the more standard values.**

Added to page 10

- **It appears that the same data were used to do exploratory and confirmatory analyses of the measurement model. The measurement model should be specified prior to the analysis and tested within the SEM framework. Using the data to determine the measurement model and then testing it within the SEM framework does not confirm the measurement model. A rationale for the measurement model that is not data driven is needed.**

This is an important issue, thus we think it deserves an extended (hopefully not “long-winded,” answer). We have tried to clarify and better present our thinking in our revisions of page 14 which present the rationale for our two measurement models (see a depiction in Figure 2). The rationale is rather simple, really, namely that the set of Health Beliefs *scales* load on a latent construct together, and similarly, that the Knowledge *scales* do also.

Additionally, it is important to note that the measurement models depicted in Figure 2 were specified prior to the analysis. A couple of modifications were made to our original conception of the measurement models (e.g., a decision to drop the Susceptibility scale as an indicator of the Health Beliefs construct due to low/non-significant factor loadings and the cross-loadings of the Benefits scale), but these are not of the type to change the essential nature of the latent constructs. Of course, we are curious to see how closely other studies replicate our models, but this is the case with any research study, regardless of the analytic procedures used.

As a sideline, and not directly related to the SEM measurement models, we wanted to validate the proposed dimensionality of the *item level* responses, to see if they had their primary factor loadings on the intended factors – this EFA procedure is what we describe on page 10. It does not seem inconsistent with the later specification and testing of a measurement model which uses scale scores as indicators, but merely a good precaution to determine whether those scale scores themselves have good measurement qualities. Please note that we had very specific *a priori* expectations about what the dimensionality of the item pool should be, and which items should load most heavily on which latent constructs. The EFA procedure that we used supported those *a priori* expectations.

In fact, as we indicated in the manuscript (p. 10), experienced statisticians such as Stan Mulaik (see, for example, Mulaik & Millsap, 2000) are now recommending that unrestricted factor analysis (essentially, EFA) be performed, preliminary to tests of measurement and structural models, in order to address this very issue.

- **The measurement model is problematic for the exercise condition as benefits of exercise loads on more than one factor. The claim that the constructs are largely unrelated is not**

supported (pg 14, line 5) when two of the constructs share a measured variable. This means that motivation and general knowledge are not unique concepts as specified by the model. Theoretically this path needs to be justified. Also, given the low loading of benefits on motivation, it is not clear why it was retained in the model.

Again, we think you raise an important point, and so treat it at length.

The Benefits of Exercise scale, used as an indicator, in the exercise measurement model, has statistically significant loadings on two latent constructs: the proposed Health Beliefs construct ($\lambda=.38$), but also the General Knowledge construct ($\lambda=.47$). The reviewer comment raises the issue of whether this creates a problem with the construct validity of the Knowledge and Health Beliefs constructs. We went back and looked carefully at our analysis results, and found that a similar cross-loading (although of smaller magnitude) should have been included for the Benefits of Calcium scale in the Calcium measurement model. (The figures and results have been revised accordingly).

We believe, in fact, that this result highlights the benefits of using a SEM approach for the analysis. If we had simply summed or averaged the scale scores to create measured composite Knowledge and Health Beliefs variables, the observed correlation between them would indeed have been a confounded estimate of the true relationship, due to the cross-loading of the Benefits scales. However, the SEM procedure allows us to isolate the two sources of variance that are driving responses to this indicator, so that each construct reflects what it is intended to, without being contaminated by the other construct. As can be observed if one looks at Figure 2, the Health Beliefs and the General Knowledge constructs are not strongly related in either measurement model. Specifically, they correlate .28 ($p < .05$) in the Calcium model and .00 (ns) in the Exercise model. These low magnitude relationships by definition tells us that Health Beliefs and Knowledge are indeed clearly separable and unique concepts.

Importantly, the fact that an indicator loads on both constructs does not mean that the constructs themselves are contaminated. (You might want to consider the analogous example of a typical EFA in which items tend to have stronger loadings on one construct than others, but rarely have loadings that are exactly zero on the other constructs. We do not typically throw out either the items or the constructs when this occurs, unless the cross-loadings or factor covariances are quite large). Relatedly, Van Prooijen & Van Der Kloot (2001) have an interesting discussion of the legitimacy of including some cross-loadings in CFA models, and note that this is often necessary or desirable because many psychological measures do not strictly adhere to a simple structure. (Also, see Mulaik & Millsap, 2000, on the issue of simple structure).

We believe that the theoretical explanation for the secondary cross-loading of Benefits on both Health Beliefs and Knowledge is that facts which are informative (“knowledge”) may also have motivational effects (“health beliefs”). People are often intrinsically motivated to perform activities that they believe will result in positive (beneficial) outcomes, especially when those outcomes relate to higher level goals such as maintaining good health. Thus, because theory and past research have included Benefits as an important constituent of Health Beliefs, and because its inclusion does not hurt the specification of either latent construct, we think it is preferable to

keep this indicator in the model. We also note a similar cross-loading for the Barriers indicator in the Exercise measurement model. A similar logic would apply here.

- **Please clearly specify what the measured variables for susceptibility are, not simply item numbers.**

The measured variables for Susceptibility are the 6 items from the Kim et al. (1991b) OHBS scale, developed by them to measure susceptibility (see p. 8). When we included the Susceptibility construct in the test of the mediator model (see Figure 4), we created 3 item parcel indicators from the 6 items, consisting of the sums of items 1&2 (parcel 1), items 3&4 (parcel 2), and items 5&6 (parcel 3). This is a fairly standard practice.

- **It seems to me that a critical path is missing from the model; one from group assignment to Daily Calcium (or Exercise) at time 1. The effect of the intervention at time 2 needs to be tested controlling for any differences on the outcome at time 1. Even if the differences at time 1 are not significant this path should be in the model. This would allow you to establish that the path from group assignment to time 2 does not disappear like the path from group assignment to time 3 when controlling for initial differences.**

This particular path was not included because it would run counter to the passage of time. Namely, participants did not receive their DXA information (the group assignment variable) until after the first set of measures was collected. Thus including this path would imply that participants' level of calcium (or exercise) was influenced by the DXA procedure which took place *after* they had already exercised or taken calcium.

Including a path from Daily Calcium at T1 to Daily Calcium at T2 (as we did) accomplishes the goal that you describe in your comment.

- **I have concerns about the conclusions drawn around mediation. The analyses as described do not support mediation as critical paths are non-significant. Susceptibility at time 1 should be included in the model to establish that group assignment was predictive of the change in susceptibility, not just time 2 values. In addition, the full test of mediation was not conducted (see Baron & Kenny 1986 and McKinnon et al 2002) so no conclusions can be drawn. Given this and the lack of a theoretical justification for mediation, the authors may want to consider dropping the mediated model from the paper.**

There is serious debate about the best way to test for mediation in SEM models (see, for example, James, Mulaik & Brett, 2006 for arguments against the Baron & Kenny, 1986 approach). Regardless, our results do not make a strong case for mediation. One of the relevant mediational paths is only marginally significant (although the other is significant). Additionally, if mediation is present, it is clearly partial rather than full mediation. However, the relationship makes theoretical sense, and we thought it beneficial to include the result in the hopes that it will be tested in other studies in the future. If requested, this result could be dropped.

- **On the figures, please include paths and coefficients for the correlations among the exogenous variables rather than the note.**

This is now done in the relevant figures (Figures 3-5).

- **Please indicate what * and + mean on the figures themselves.**
- **Define all abbreviations used in the figures in a note on *Figure Legends* page (i.e., *, +, ns)**

This information is now indicated in the figure notes, as requested.

References

- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Boston: Houghton Mifflin Co.
- Mulaik, S. A., & Millsap, R. E. (2000). Doing the four-step right. *Structural Equation Modeling*, 7, 36-73.
- Van Prooijen, J.-W., & Van Der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*, 61, 777-792.