

Feb 09, 2009

RE: NRES-D-08-00245, titled "Use of the Concordance Correlation Coefficient when Examining Agreement in Dyadic Research"

Dear Dr. Quinn,

Recently you sent your manuscript for consideration by Nursing Research. The review of your manuscript is complete and the comments of the reviewers are listed below along with a checklist for style.

Your reviewers made helpful recommendations. When you address them, please give special attention to showing results for the three statistical tests that you discuss and explain why the kappa statistic was not selected for use. You may revise your manuscript based on the reviews and resubmit the manuscript to Nursing Research for further consideration. I will verify that this has been done upon receipt of the revised manuscript.

Please include with your revised submission an itemized, point-by-point response with page and line numbers to the comments of the reviewers. The revisions should be completed by Apr 10, 2009 to avoid being considered as a new submission.

To submit a revision, go to <http://nres.edmgr.com/> and log in as an Author. You will see a menu item called "Submission Needing Revision." Please click on this item to obtain your submission record and begin the revision process.

With Kind Regards,

Molly C. Dougherty, PhD, RN, FAAN  
Editor  
Nursing Research

Reviewer Comments:

Reviewer #1: General Comments- The primary purpose of this paper is to evaluate the use of the concordance correlation coefficient (CCC) to compare agreement of symptoms between patients with HF and their family members. This paper makes a contribution to our understanding of inter-rater reliability and correlations; common statistical tests that are not addressed frequently in the literature. One complaint from practitioners is the difficulty in comprehending the statistical analyses in research articles. This paper could serve a broad audience of students, researchers, and practitioners.

Background Literature- Acceptable.

Theoretical Framework- The research questions are based on an examination of three statistical formulas for correlation and an analysis of the agreement of symptoms of HF between patients and family members.

Research Design and Methods- This is a secondary analysis of a study of patients with HF and their family members. Symptom data were collected from both patients

and their family members.

The choice of a statistical test hinges on the research question. In this case, why do the researchers want to evaluate congruence of symptoms between patients and family? Are they interested in inter-rater reliability? Do they want to know if there are statistical differences in ratings on which to make clinical decisions? Curiously the kappa statistic (Landis & Koch, 1977) was not used. If the goal is to evaluate inter-rater reliability-this more commonly used statistic would control for expected agreement between symptoms that are frequently reported such as fatigue and dyspnea. The Phi statistic has also been used to evaluate congruence between raters.

The authors state that "the ICC and PCC are used to examine if measures vary in the same direction." P4, L2-3. However, the PCC measures the strength of an association, whether it is in the same direction or not (i.e. positive or negative). So the case for using PCC or CCC remains unclear. The authors go on to state that "the CCC was developed and tested in the biopharmaceutical industry to assess agreement between methods of measurement" (P4, L8-9). This implies that it is used in the case of multiple measures. For example, there are various methods to test whether a drug is "effective" including physiologic measures, serum drug levels, or biomarkers. Correlations would yield some information about the relationships of the measures but that is not the same as examining congruence between two raters administering the same measure. One would hope to achieve perfect inter-rater reliability but would not expect perfect correlations between different measures.

Data Analyses- Congruence between symptoms was measured in three ways; with the standard Pearson correlation coefficient (PCC), intraclass correlation coefficient (ICC), and the CCC. Each of these tests are detailed in the Methods section. Please see specific comments and recommendations below.

Results- No symptoms had good agreement between patient and family. This seems like a potentially serious problem particularly if the family member is serving as a proxy for the patient. Please address this.

Discussion- See above. The discussion of the coefficients is thorough. When looking at the findings in Table 3. It is apparent that the symptoms with the least agreement are probably the symptoms that are most frequently reported hence (as with the Kappa statistic) the findings are not unexpected. Could you address this?

Organization and Style- Well organized and well written.

Summary- The topic is of broad interest. Addressing statistical issues is helpful and HF is a serious problem and growing.

Specific recommendations:

P6, L9-10-Agreement between 2 observers is usually evaluated via the PCC or ICC- as above (did you consider using the Kappa or Phi statistic?).

Please state why the examination of congruence in symptoms between patient and family is important. (P9, L14-16 describes what you are interested in-absolute agreement-but not why).

P3, L 20- Recommend Pearson's r coefficient (lower case) or simply Pearson's

correlation coefficient.

P8, L16-17-If the CCC requires that score differences between observers be small, won't an 11-point scale result in potentially large differences?

Table 1. Example 3 seems unnecessary as it is well understood that perfect agreement will result in a perfect correlation ( $r=1$ ).

Table 3. The mean scores for the symptoms would be of interest.

Reviewer #2: The manuscript is well written and provides a valuable tool to researchers studying symptom perceptions among family members. Three approaches to rating the relationship between patients and family members are presented in a clear and concise manner.

There are two issues, if addressed, will help strengthen this manuscript. The first is discussion comparing the ICC(A,1) and the CCC. The authors state the CCC is easier to calculate (which I agree is true) than the ICC(A,1) so they refer to the ICC(A,1) as the CCC. Are these numerically the same? If so, that should be stated and shown or at least a reference should be provided. If they are not the same, then the sentence should be changed.

The second issue that should be addressed is the discussion of the results. It would be useful to show the results of all three methods, and then to compare and contrast the differences between them.

One additional minor point - on page 7, line 11, I believe a should equal 4 not 3.

---

#### CHECKLIST FOR STYLE

TEXT -- Please use Times New Roman font with 12 point pitch.

FIGURE 1 -- Use x & y axes only (remove top and right side lines)