

University of York Department of Health Sciences

**M.Sc. in Evidence Based Practice**

**Measuring Health and Disease**

**Assessment, June 2007**

Time allowed: 60 minutes

**ANSWER TWO QUESTIONS**

Each part of a question carries equal marks.

*This is a closed book exam. Questions relate to three papers which will be given out one week in advance of the exam. Clean copies will be available at the examination itself.*

**Question 1.**

This question relates to the paper 'Muscle strength testing with one repetition maximum in the arm/shoulder for people aged 75 — test-retest reliability'.

- (a) In Figure 3 we see the results of sessions 1 and 2 in kg for the combined group, with the line of equality. What is the line of equality? Why is used here, rather than a regression line?
- (b) In Figure 3, the authors quote a correlation coefficient  $r = 0.97$ ,  $P < 0.0001$ . What null hypothesis is being tested and what is its value here? What feature of the study design is likely to inflate the value of this correlation coefficient? What feature of the results would be missed by the use of a correlation coefficient?
- (c) Figure 4 shows 95% limits of agreement analysis comparing measurements in the first and second sessions for subjects in group 1. What are 95% limits of agreement? What is the purpose of this plot and what does it show here? What should we conclude about the limits of agreement estimated?

**Question 2.**

This question relates to the paper 'Development of a pictorial scale of pain intensity for patients with communication impairments: initial validation in a general population'.

- (a) In Table 3, the authors present the test-retest reliability of the visual analogue scale and numeric rating scale by intraclass correlation coefficients. What is meant by 'test-retest reliability'? Why are intraclass correlation coefficients used to estimate it for these two scales?
- (b) In Table 3, the authors present the test-retest reliability of the scale of pain intensity evaluated by weighted kappa coefficients. What is a weighted kappa coefficient and how would we interpret the kappa values presented here? In the Methods, these are described as quadratic-weighted kappa statistics. What does this mean and why is it useful information?
- (c) In Table 3, the authors present 95% confidence intervals for the weighted kappa and the intraclass correlation coefficients. Why are several of these obviously wrong?

### Question 3.

This question relates to the paper 'Validity of Hamilton Depression Inventory in Parkinson's Disease'.

- (a) Table 2 shows sensitivity, specificity, positive predictive value, and negative predictive value for different cut-off points and different target diagnoses. What are 'sensitivity', 'specificity', 'positive predictive value', and 'negative predictive value'? For each scale, as we go down the table, sensitivity decreases and specificity increases. Why is this?
- (b) Figures 1 and 2 show ROC curves. What is an ROC curve and why might it be useful? Why are the areas under the ROC curves given and how can they be interpreted?
- (c) The authors describe this as a study of the concurrent validity of the HDI-17. What is meant by 'concurrent validity'? In this study, would this be best described as an aspect of criterion validity or of construct validity, and why?