University of York Department of Health Sciences

M.Sc. in Evidence Based Practice

Measurement in Health and Disease

Specimen Assessment, June 2005

Time allowed: 60 minutes

ANSWER TWO QUESTIONS

Each part of a question carries equal marks.

Question 1.

The following is the abstract from a recent paper:

Background. In subjects with type 1 diabetes, persisting elevations of fetal hemoglobin (HbF) have been demonstrated. This study evaluated whether HbF levels typically seen in type 1 diabetes (up to 3%) interfere with glycohemoglobin determinations using a common immunologic method (DCA 2000(TM)).

Methods. HbA(1c) was measured by high-performance liquid chromatography (HPLC) using a Diamat(TM) analyzer in 90 type 1 diabetics with parallel determinations of HbF. Results were compared with HbA(1c) concentrations obtained using DCA 2000(TM).

Results. Reproducibility was good for both methods with coefficients of variation <5% and correlation between the two methods was good (r^2 = 0.939, p <0.0001). Mean difference between the two methods was small (0.007%). Limits of agreement varied between -0.92% and +0.93% and constant bias (intercept: 0.73, 95% CI 0.28-1.18) as well as a proportional bias (slope: 0.92, 95% CI 0.87-0.97) were detected. At low concentrations of HbF, the DCA 2000(TM) immunologic method tended to underestimate and at higher concentrations tended to overestimate HbA(1c) when compared with Diamat(TM).

Conclusions. DCA 2000(TM) allowed measurements of HbA(1c) rapidly and with precision adequate for clinical purposes. However, agreement with Diamat(TM) results was comparatively weak with both constant as well as proportional biases. The 95% limits of agreement between Diamat(TM) and DCA 2000(TM) fell within a range that significantly limited traceability between these two methods; therefore, the two methods should not be used interchangeably. Small but persistent elevations of HbF concentrations were identified as a significant cofactor, which may be relevant for limited traceability between the two methods.

(Diem P, Walchli M, Mullis PE, Marti U. Agreement between HbA(lc) measured by DCA 2000 and by HPLC: Effects of fetal hemoglobin concentrations. *Archives of Medical Research* 2004; **35**: 145-149.)

a) What is meant by ‘coefficients of variation <5%’?

b) What are the disadvantages of correlation as a measure of agreement?

c) What is meant by ‘limits of agreement’?

d) Why would we expect a positive intercept and slope to be less than 1.0?

e) ‘At low concentrations of HbF, the DCA 2000(TM) immunologic method tended to underestimate and at higher concentrations tended to overestimate HbA(1c) when compared with Diamat(TM).’ Is this a valid conclusion from the regression analysis?

Question 2.

The following is the abstract from a recent paper:

BACKGROUND: Left ventricular function is the most important determinant of prognosis following myocardial infarction.

METHODS: A prospective analytical cohort study of 33 cardiac care unit survivors of acute myocardial infarction was performed to assess the accuracy, reproducibility and observer variation of the bedside Valsalva response in predicting an ejection fraction (EF) less than 40%.

RESULTS: Agreement between physicians for the clinical Valsalva response was excellent (kappa coefficient = 0.75), as was the comparison between physicians' clinical response to the Finapres (Ohmeda, USA) hemodynamic response (weighted kappa = 0.85). The EF was significantly higher among patients with a normal Valsalva response (56.9%) than in patients with either an absent overshoot (48.4%) or a square wave (28.3%) response ($P<0.001$). Physicians were very accurate at estimating whether the EF was greater or less than 40%. In 40 of 66 situations, the clinicians were confident, based on the clinical examination and the Valsalva response, that the EF was either greater or less than 40%. In these situations, agreement (95.0%) and kappa (0.89) were both excellent. When all patients were considered, the degree of agreement (90.8%) and kappa (0.80) diminished slightly. A square wave response had poor sensitivity (37.5%) but excellent specificity (92.7%), whereas any abnormal response had excellent sensitivity (91.7%) but poor specificity (54.8%).

CONCLUSION: Compared with many other aspects of the clinical examination, the bedside Valsalva manoeuvre has acceptable degrees of interobserver variability. A normal response tends to rule in an EF of greater than 40%, whereas a square wave response rules out an EF of greater than 40%. This simple bedside manoeuvre may be useful in predicting low EF following acute myocardial infarction.

(Massel D, Marchiori G. Precision and accuracy of the bedside examination in detecting an ejection fraction of less than 40% following acute myocardial infarction. *Canadian Journal of Cardiology* 2004; **20**: 411-416.)

a) What was meant by ‘kappa coefficient = 0.75’?

b) What was meant by ‘weighted kappa = 0.85’? Under what circumstances would we use weighted kappa?

c) What was meant by ‘sensitivity (37.5%)’? What does it tell us about using square wave as a test for low ejection fraction?

d) What was meant by ‘specificity (92.7%)’? What does it tell us about using square wave as a test for low ejection fraction?

e) Why did sensitivity increase but specificity decrease when any abnormal response was considered, compared to a square wave response only?

Question 3.

The following is the abstract from a recent paper:

Background: Over the last 40 years the study of human attitudes toward death has attracted much scientific interest and a significant amount of research has been carried out in the English-speaking world. However, among Spanish-speaking researchers the subject has been practically ignored, as has the issue of related psychometric instruments. The aim of this study was to develop the Death Anxiety Inventory (DAI) and thus provide a valid and reliable assessment instrument for measuring death anxiety among Spanish-speaking subjects.

Methods: This study examined the psychometric properties of the DAI. The DAI is a self-administered questionnaire of 20 items that can be used in either a dichotomous true/false format or on a six-point Likert scale. The properties of both scales were investigated by means of six empirical studies and several samples.

Results: The scale has an alpha coefficient of internal consistency of 0.90 and test-retest correlation, at 4 weeks, of 0.94. The correlation with Templer's Death Anxiety Scale was 0.79. Factor analysis of the DAI identified five significant factors. Taken together these factors explained 54.60% of the total variance and were labelled as: (1) Externally generated death anxiety, (2) Meaning and acceptance of death, (3) Thoughts about death, (4) Life after death, and (5) Brevity of life. The English form of the DAI is also presented in the study in order to enable cross-cultural comparisons to be made.

Conclusion: Results of this study suggest that the DAI has adequate psychometric properties that make it a valid and reliable instrument to assess death anxiety in Spanish-speaking individuals.

(Tomas-Sabado J, Gomez-Benito J. Construction and validation of the death anxiety inventory (DAI). *European Journal of Psychological Assessment* 2005; **21**: 108-114.)

a) What is a test-retest correlation at 4 weeks and what does a value of 0.94 tell us?

b) What is the ‘alpha coefficient of internal consistency’ (Cronbach’s alpha) and what does a value for alpha of 0.90 tell us?

c) What is factor analysis and how do is practitioners decide on the names and meaning of their factors?

d) What does it mean when the authors say that ‘these factors explained 54.60% of the total variance’?

e) Why do the authors give us the correlation with Templer's Death Anxiety Scale?