International Journal of Cardiac Imaging **13:** 173, 1997.

Editorial comment

In response to the article by H.J. Muntinga et al., Normal values and reproducibility of left ventricular filling parameters by radionuclide angiography (see pp. 165–171)

When is a visually or mathematically diagnosed scan abnormality clinically important?

This question is frequently addressed as more and more sensitive and sophisticated diagnostic tools become available.

The first step in setting up a diagnostic test is to evaluate its reliability in terms of sensitivity, specificity and reproducibility under a set number of conditions. In a diagnostic imaging department, this is usually achieved by repeat imaging and analyses.

In this respect the authors have satisfied the basic requirement necessary to introduce the test. The next question, of course, is the test being utilised on a regular basis? If so, is there an expectation to supplement the results of this paper with additional data (preferably a combination of diseased and normal subjects) so that validity of the repeated measurements will be tested to determine the specificity, sensitivity and accuracy of the method. Without this additional data the reproducibility figures derived from repeated measurements in small studies can surely only be of interest and importance as a quality control procedure. It is also well recognised that the normality in parameters derived by radionuclide cardiac angiography are geographically and population biased so that multicentre trials to establish normal values are essential. The pooling of data using meta-analyses has limitations and inherent deficiencies so that this choice of analysis should be avoided where possible.

Presenting p-values alone can lead to their being given more merit than they deserve. There is a tendency to equate statistical significance with medical importance or biological relevance. Small differences of no real interest can be statistically significant with large sample sizes, whereas clinically important effects may be statistically non-significant only because the number of subsets studied was small.

This phenomenon is clearly evident in this study where a number of p-values are shown to indicate significant differences and where the correlation coefficients demonstrate large spread of individual data. Including data on diseased subjects will test the reliability of the hypotheses. The estimation and use of confidence intervals would have been more appropriate.

In larger studies on impaired left ventricular function and diastolic filling Bonow et al. and Pace et al. have concluded that 'the differences between the diastolic filling characteristics of patients with coronary artery disease without prior infarction and those of normal volunteers could not be explained on the basis of any differences in age, in heart rate or in left ventricular end-diastolic size. . .'. This finding is not supported by this study and may be a result of the geographic, population sample and methodological bias.

The evaluation of reproducibility is surely of much importance provided that studies are designed to go beyond laboratory based quality control and include data to test hypotheses on a broader base, i.e. sensitivity, specificity and accuracy.

References

Bonow RO et al. Impaired left ventricular diastolic filling in patients with coronary artery disease: assessment with radionuclide angiography. Circulation 1981; 64: 315–23.

Pace L et al. Diagnosis of coronary artery disease by radionuclide angiography: effect of combining indices of left ventricular function. J Nucl Med 1989; 30: 1966–71.

Dr. Daniel den Hoed Clinic M. Pillay
University Hospital Rotterdam
Fax 010-4863357