

Interobserver variability in interpretation of coronary arteriograms

Harvey G. Kemp, Jr., M.D.

New York, New York

Several recent studies suggest that there may be more interobserver and intraobserver variability in the interpretation of coronary arteriograms than had been previously thought to exist.¹⁻³ Because of these data, a study was initiated utilizing data from the Coronary Artery Surgical Study supported by the National Heart, Lung and Blood Institute. This is a multicenter study of the effect of coronary artery bypass surgery on the natural history of coronary artery disease. The centers participating have entered the interpretation of coronary arteriography on consecutive patients studied into a data bank at the Coordinating Center at the University of Washington in Seattle. Three of the centers have served as quality control centers reinterpreting an aliquot of films from each of the other centers. *These data form the basis of this report.*

A pilot study was performed by the three quality control centers. Thirty films were chosen by the Coordinating Center excluding films that were considered to be normal. These 30 films were divided into groups of ten films and were sent to each quality control site for a consensus interpretation. The films were then exchanged until each site had read each group of ten; and then the process was repeated for a second round of interpretation of the same films. In this manner, both interobserver and

intraobserver variability could be assessed.

We initially studied agreement on the number of vessels involved. A vessel was considered significantly stenosed if it contained a lesion of 70% or greater in any major coronary vessel with the exception of the left main coronary artery, which was considered significantly involved if it contained a lesion of 50% or greater. Intraobserver variability (a comparison of reading one with reading two by the same center) varied from 60% to 77% agreement. The combined data yielded an agreement of 68%. The 32% nonagreement was equally divided between underreading and overreading on the second reading. Interobserver variability showed agreement 61% of the time with two sites underreading and overreading approximately the same and a third site consistently underreading when compared to the other two. It would appear from these data that there is approximately a 60% chance for two different centers to agree on the number of vessels involved if a careful consensus reading is performed.

A second analysis of these data was performed plotting a regression line for percent stenosis for the different segments of the coronary circulation; *r* values for these analyses for within-site readings of the proximal left anterior descending varied from .78 to .94. Between-site interpretations varied from .73 to .82. Analysis of variability of interpretation of lesions in the mid-left anterior descending was almost as good with *r* values for within-site interpretation varying from .80 to .85, and between-site interpretations from .77 to .82. Analysis of the proximal right coronary artery was quite similar: variability for within-site interpretation ranged from .84 to .94 and between-site from

.74 to .85. On the other hand, variability increased considerably when the proximal left circumflex was studied. Within-site variability ranged from .59 to .85 and between-site from .65 to .70.

Relatively few patients in this study had disease in the left main coronary artery. Only seven cases were available for analysis. Two sites showed perfect agreement on two successive readings, whereas a third site read two cases as having a 50% stenosis on the first reading and none on a second reading. Both of these cases were read as either normal or nonsignificant by the other two sites. Because so little data were available on the interpretation of left main lesions and because of the relative importance of correct interpretation of disease in the left main, a second study was conducted utilizing 50 films in which a 50% or more stenosis was called by either the initial participating site or a quality control center. These films were collected and reinterpreted for a third time by a panel of three to ten experienced angiographers. Comparing the panel interpretation with the quality control interpretation yielded the highest level of agreement, which was approximately 66%. Calcified lesions of the left main coronary artery and diffuse disease of the left main coronary artery offered particular problems in interpretation and increased interobserver variability.

Analyses were also performed looking at smaller segments of the coronary circulation such as the posterior descending artery, distal left anterior descending artery, distal marginal branches, and diagonal branches. The variability in interpreting lesions in these smaller branches was so great as to make regression analysis meaningless.

These studies suggest that lesions in the proximal portions of the major cor-

onary trunks can be read with acceptable interobserver and intraobserver variability, whereas lesions in the more distal coronary circulations generally cannot. The left main coronary artery presents some particular problems and agreement can be achieved only two thirds of the time. Studies that involve analysis of coronary disease based on the number of vessels involved have a distinct limitation because of interob-

server and intraobserver variability in making this quantification.

References

1. Detre KM, Wright E, Murphy ML, et al: Observer agreement in evaluating coronary angiograms. *Circulation* **52**: 979-986, 1975.
2. Zir LM, Miller SW, Dinsmore RE, et al: Interobserver variability in coronary angiography. *Circulation* **53**: 627-632, 1976.
3. DeRouen TA, Murray JA, Owen W: Variability in the analysis of coronary arteriograms. *Circulation* **55**: 324-328, 1977.