Factors Affecting Interpretative Accuracy: How Can We Reduce Errors?¹

In the accompanying article, “Effect of Shift, Schedule, and Volume on Interpretative Accuracy: A Retrospective Analysis of 2.9 Million Radiologic Examinations,” Hanna et al (1) demonstrate that long shifts and high examination volume lead to increased interpretative discrepancies between preliminary and secondary readers. This analysis provides insight into reasons for diagnostic errors and adds to the body of knowledge that should help us to improve our diagnostic accuracy.

There are certainly other factors that affect our accuracy, and one that possibly needs more attention is the advanced training and experience of subspecialty-trained radiologists. The authors of the accompanying article mention that some of the preliminary and subsequent readers have subspecialty training, but there is no mention that they concentrated on examinations that matched their subspecialties. Authors of a number of articles have analyzed the effect of such subspecialization and have demonstrated significant improvement in diagnostic accuracy in various areas.

Sickles et al (2) examined the effect on both screening and diagnostic mammography and found that the subspecialists detected significantly more cancers, recommended more biopsies, and had lower recall rates than did the generalists. Mullerad et al (3) analyzed the predictive value of various techniques in endorectal magnetic resonance imaging for prostate cancer. They found that in the hands of genitourinary radiologists, the diagnosis of extracapsular extension was substantially improved.

In pediatric radiology, Eakins et al (4) examined patients transferred to a tertiary pediatric hospital and found major discrepancies in radiologic interpretations when compared with those from the referring institution. On the other hand, Franken et al (5) analyzed the effect on pediatric radiology at a rural hospital and showed that general radiologists’ interpretations were equal to those of the pediatric radiologists for this population who presented mostly for acute care of pneumonia and fractures.

Another important factor to consider is double reading of examinations. There are multiple variations on this subject such as residents or fellows reading with attending radiologists; two radiologists reading together, such as in mammography; and preliminary followed by subsequent readers. In Norway, Husby et al (6) explored double reading and found it common for positron emission tomography/computed tomography, mammography, and, of course, in resident training. However, there is no doubt that double reading has a negative financial effect.

In a review of the literature, Dinnes et al (7) described the advantages of double readings in screening mammography and found improved accuracy in cancer detection and a reduction in the recall rate. However, Anderson et al (8) examined screening mammography and found that double readings increased sensitivity but showed a slight decrease in specificity, with more unnecessary recalls.

In a relatively small study performed to examine the detection of lung cancer on the basis of chest radiography, Quekel et al (9) found increased sensitivity and slightly reduced specificity with double readings and increased sensitivity without reduced specificity with dual readings. Unlike dual readings with two radiologists side by side, the double readings were done separately.

General radiologists have provided teleradiology services extensively, especially at night and over weekends. These services are often the only interpretation for these examinations, but sometimes there is a subsequent second reading the next day. Many
academic departments use preliminary readings by residents at night followed by interpretations by attending physicians the next day. Authors of several studies have examined diagnostic error rates by comparing these different approaches. In general, overall error rates are quite low, and most errors had minimal effect on patient care. Residents actually performed better than general radiologists providing coverage at night. Even if infrequent, the rare significant errors that do occur can affect care and lead to bad outcomes (10,11).

One can conclude that, for fairly simple outpatient care and possibly overnight emergency care and urgent inpatient care, general radiologists or even residents overnight provide good care. The error rates by residents were often similar to interobserver differences among subspecialists for overnight interpretations. The real advantages for subspecialists may be for advanced procedures, techniques, and complicated diseases, especially in the inpatient setting. There is little doubt that dedicated mammographers, pediatric radiologists, and neuroradiologists raise the bar substantially for diagnostic accuracy. If we are striving for reduction in medical errors, using more subspecialists makes sense, especially in selected settings.

Double reading and dual reading have the advantage of four eyes and two brains, but the negative financial implications keep this approach from being practical in most situations. Mammography may well be an exception. In addition, patients referred for specialized care benefit from subsequent readings by subspecialized radiologists.

If medicine moves more toward physician payment based on quality or outcomes, the financial equation could shift. Under those circumstances, it might be better financially as well as for our patient care if we include double readings and readings by subspecialists more often.

What conclusions can we draw about the relative significance of long shifts and heavy volume versus subspecialty interpretations and double readings? They all make a difference, and, in a future world where we are rewarded more for quality and outcomes, we will probably start to address them all.

Disclosures of Conflicts of Interest: R.L.A. disclosed no relevant relationships.

References