

## FALSE POSITIVE AND FALSE NEGATIVE RESULTS OF THYROID CYTOPATHOLOGY REMAIN A CHALLENGE

Wang CC, Friedman L, Kennedy GC, Wang H, Kebebew E, Steward DL, Zeiger MA, Westra WH, Wang Y, Khanafshar E, Fellegara G, Rosa J, Livolsi V, Layman RB. **A large multicenter correlation study of thyroid nodule cytopathology and histopathology.** *Thyroid* 2011;21:243-51. Epub December 29, 2010.

### SUMMARY ●●●●●●●●●●●●●●●●●●●●

#### BACKGROUND

The authors aimed to : (1) review the diagnostic performance of thyroid fine-needle aspiration (FNA) biopsy in a prospective study, (2) perform a meta-review of large FNA studies from centers in the United States published between 2002 and 2010, and (3) assess the concordance of diagnosis when experts review surgical pathology.

#### METHODS

For the prospective study, FNA specimens and clinical data were collected from 16 U.S. community-based clinics, 3 U.S. academic centers, and 2 non-U.S. academic sites from August 2008 through January 2010. The diagnostic categories were benign, malignant, indeterminate, and nondiagnostic. The indeterminate category lumped together the three separate categories of the Bethesda system: atypia of undetermined significance or follicular lesion of undetermined significance; follicular (or Hürthle) neoplasm or suspicious for follicular (or Hürthle) neoplasm; and suspicious for malignancy. Eleven studies were selected for the meta-review; the total number of patients was 17,059, of whom 8937 had surgical specimens.

#### RESULTS

In the prospective study, 753 FNA biopsies were performed; 80% were categorized as benign, 8%

indeterminate, 7% malignant, and 5% nondiagnostic. There were 112 surgical specimens; the malignancy rate in relation to the FNA category was 11% of the benign, 34% of the indeterminate, 98% of the malignant, and none of the nondiagnostic category. Based on these data, the authors calculated that FNA biopsy had a sensitivity of 95%, a positive predictive value of 68%, and a negative predictive value of 89%. These results were similar to the diagnostic performance of FNA biopsy in the meta-review.

In the meta-review using the Bethesda system of classifying FNA biopsies, the malignancy rate was 12% in the benign category, 16% in the atypia, 25% in the follicular lesion, 62% in the suspicious, 97% in the malignant, and 12% in the nondiagnostic category.

In a review of surgical material by two expert pathologists, there was 11% disagreement with regard to categorizing benign versus malignant between diagnoses by experts and local pathologists and 8% disagreement between the two experts. When the experts conferred about their disagreements, the difference was reduced to 3%.

#### CONCLUSIONS

False positive and false negative results of thyroid cytopathology remain a challenge. The surgical pathology was benign in two thirds of the indeterminate FNA cytology category.

### COMMENTARY ●●●●●●●●●●●●●●●●●●●●

This report raises several issues. First, it should be noted that this paper emanated from the Veracyte company, which markets a thyroid FNA molecular diagnosis method that is an analysis of 142 genes.

It is surprising that, in the prospective study, only 8% of lesions were classified as indeterminate, whereas in the meta-review, the indeterminate category was found in 10% to 26% of reports. In addition, by ignoring the breakdown of the indeterminate category into its three subtypes that had progressively more possibility

of malignancy going from atypia to suspicious for carcinoma, the authors lost the benefit of the Bethesda classification. This is shown in the results of the meta-review: a relatively low incidence of malignancy in the atypia category (16%) and a high incidence in the category suspicious for malignancy (62%). These figures must be tempered by the realization that the managing clinician probably used other factors, such as size, appearance on the ultrasound, family history, patient's concern, etc. in the decision as to whether to send a patient for surgical removal of the nodule. This is emphasized by the fact that 11% to 12% of the benign cytology category resulted in diagnoses of malignant lesions at surgical pathological review, although a few were microcarcinomas.

When the clinical factors as noted above do not suggest malignancy, I generally do not recommend surgery for the atypia category because I believe that the

likelihood of malignancy does not justify the potential complications of surgery. It should be noted that the authors used the indeterminate FNA classification as a positive diagnosis of cancer in their calculations of predictive values. Of course this increases the false positive percentage.

There is a need for improvement in the diagnosis of thyroid cancer, and molecular markers may provide the progress that is needed in this area. However, these biomarkers for malignancy or benignity are unlikely to entirely replace clinical judgment. In addition, the use of FNA biomarkers depends on obtaining an appropriate sample. And of course, at present there would be a substantial financial cost for molecular markers that examine a large number of genes.

**Jerome M. Hershman, MD**