Cytologic Artifacts and Pitfalls of Thyroid Fine-Needle Aspiration Using ThinPrep®

A Comparative Retrospective Review

Alaa M. Afify, M.D.
Jing Liu, M.D.
Basim M. Al-Khafaji, M.D.

Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan.


Address for reprints: Alaa M. Afify, M.D., Department of Pathology, University of Michigan Medical School, 1500 E. Medical Center Drive, 2E322, Box 0054, Ann Arbor, MI 48109; Fax (734) 763-4095; E-mail: alaa@umich.edu

The brand names ThinPrep®, Cytolyt®, and PreserveCyt® used throughout this article are registered trademarks of the Cytyc Corporation, Boxborough, MA.

Received July 20, 2000; revision received January 10, 2001; accepted January 17, 2001.

BACKGROUND. The ThinPrep® Processor has gained popularity as a collection and preparation technique for fine-needle aspiration biopsy (FNAB). Specific cytologic criteria to evaluate ThinPrep preparation (TP) may differ from those of conventional preparation (CP). The authors retrospectively reviewed the quality, cytologic features, and pitfalls of TP versus CP in thyroid FNABs and addressed the cytomorphologic criteria used to evaluate TP specimens.

METHODS. Thyroid FNABs received between January 1996–July 1999 were identified from the computer files of the Department of Pathology, University of Michigan (Ann Arbor, MI). Histologic correlation and clinical follow-up were reviewed. The cytology slides were reevaluated for cellularity, cellular preservation, artifacts, background material, architectural integrity, cytoplasmic details, and nuclear details by two observers.

RESULTS. Of the 209 thyroid FNABs performed during the study period, TP and CP prepared 127 and 82 cases respectively. Histologic correlation was available in 68 (33%) cases (32 TP and 36 CP). Overall sensitivity was 80% and specificity was 98%. The sensitivity of CP versus TP was 87% and 70%, respectively. Thyroid FNABs prepared by TP, as compared with CP, were characterized by the following: The TP slide 1) allowed assessment of the overall specimen cellularity but not individual passes of an FNAB, 2) contained only “hard” colloid that appeared dense, markedly fragmented, or in droplets, 3) showed crowded, tight, tissue clusters with loss of cellular preservation, especially in the larger aggregates, 4) demonstrated more cell shrinkage, 5) showed increased disruption of the cytoplasm and numerous naked nuclei, 6) occasionally gave nucleoli a more prominent appearance, and 7) was less likely to show nuclear grooves and “pseudoinclusions” in papillary carcinoma.

CONCLUSIONS. This study concluded that cytologic features used to evaluate thyroid FNABs prepared by CP may need to be modified when using TP. Awareness of the above-described findings and further studies to evaluate TP are essential to avoid potential diagnostic pitfalls.

© 2001 American Cancer Society

KEYWORDS: ThinPrep®, fine-needle aspiration, thyroid.

Fine-needle aspiration biopsy (FNAB) of the thyroid gland currently is considered the best method to assess thyroid nodules. Cytologic examination of FNAB involves several parameters; it starts with appropriate aspiration technique, followed by suitable processing methods, and concludes with microscopic evaluation by a pathologist. To retain FNAB as the most sensitive and most specific simple diagnostic test available for evaluation of thyroid nodules,¹ we need to assure continuity of the acquisition of high-quality cytologic specimens. This will depend upon choosing a preparation method that will preserve morphologic detail for an accurate diagnosis.

The ThinPrep® Processor is a slide-preparation device that re-
cently has become widely available\textsuperscript{2,3} and has gained popularity as a collection and preparation technique for both gynecologic and nongynecologic cytologic specimens. It prepares a monolayer of cells on a glass slide from cells collected in an alcohol-based preservative solution (Cytyc CytoLyt).\textsuperscript{4} Several studies have suggested that ThinPrep (TP) has an improved diagnostic sensitivity over conventional preparation (CP), especially in gynecologic\textsuperscript{5} and certain exfoliate, nongynecologic, cytologic specimens such as urine.\textsuperscript{6} However, several artifacts were observed, in particular, within specimens in which the diagnosis might have relied on architectural integrity, stromal element, or background constituents.\textsuperscript{7} Further, several of the previous studies have depended on the evaluation of specimens obtained from resected surgical pathology specimens and not from the evaluation of "real life" patient specimens, the acquisition and processing of which involves several other parameters as mentioned above.

This study aimed to retrospectively review and compare the diagnostic accuracy, quality, cytologic features and pitfalls of TP versus CP in thyroid FNABs and to evaluate the reproducibility of the cytomorphologic criteria used in the evaluation of FNABs prepared by CP versus TP.

**MATERIALS AND METHODS**

Thyroid FNABs received between January 1996 and July 1999 were retrieved from the computer files of the Pathology Department, University of Michigan Hospital (Ann Arbor, MI). Patient samples were obtained through FNAB performed by a cytopathologist, a clinician, or a radiologist who used ultrasound-guided fine-needle aspiration. Whenever samples were obtained by a cytopathologist, smears were air dried, stained with DiffQuik\textsuperscript{®} or Papanicolaou methods, and fixed in 95\% ethanol. Additionally, the syringe used for FNAB was rinsed in CytoLyt Solution and additional cellblock and/or TP preparations were prepared if deemed necessary. Samples obtained by clinicians and/or radiologists were received in CytoLyt Solution (a transport reagent). These were centrifuged, resuspended in PreservCyt Solution (a preservative solution), and processed using the TP processor according to the manufacturer’s instructions.\textsuperscript{4} Further, cellblock preparations also were made from these specimens. All TP slides were stained using the Papanicolaou procedure according to the manufacturer’s instructions. Cytologic diagnoses were classified into four major categories as follows: nondiagnostic, negative for malignancy, atypical/indeterminate, and positive for malignancy. The atypical/indeterminate category encompassed the follicular or Hürthle cell lesions, which could not be further classified by cytologic criteria into a definitively benign or malignant process. The differential diagnosis in such lesions usually is broad and may include a dominant cellular nodule in nodular hyperplasia, follicular adenoma, or follicular carcinoma. This category served its purpose in alerting clinicians that additional tissue sampling was necessary. Therefore, clinicians should correlate the FNAB diagnosis with the patient history and the clinical findings to determine the management plan.\textsuperscript{8}

As has been reported previously, the wide range of reported sensitivities and specificities reflects, in part, a difference in how atypical findings are evaluated.\textsuperscript{9,10} If an atypical result is considered positive, sensitivity will be enhanced at the expense of specificity. Conversely, if an atypical result is considered negative, specificity will be enhanced at the expense of sensitivity.\textsuperscript{11}

A specimen was considered nondiagnostic if it had an inadequate number of follicular cells, and/or poor cellular preservation, or obscuring blood. The rule, which we required for specimen adequacy on CP, was the presence of at least six or more clusters of cells, all of which had to be benign on at least two slides before we accepted a specimen as adequate.\textsuperscript{12,13} Each cluster or group had to contain at least 10 cells.\textsuperscript{14} We applied this rule to TP, and, thus, a minimum of 12 clusters of follicular cells (each with at least 10 cells) could be used for specimen adequacy. Such guidelines helped to direct us in our interpretation. However, if the quality of the smears was such that we had to consider counting the cells, then the specimen was judged as probably inadequate.

Additionally, slides were evaluated for cellular preservation (satisfactory or degenerated), staining characteristics, artifacts, background material, colloid (quality and quantity), cellularity (hypocellular, normocellular, or hypercellular), architectural integrity, cyttoplasmic details, and nuclear details. Cytologic and histologic diagnoses were correlated, when available, and the results were classified into four categories: True-negative: The absence of a malignancy was correctly diagnosed cytologically. True-positive: The presence of a malignancy was correctly diagnosed cytologically (including smears, which were read as suspicious for malignancy). False-negative: The cytologic smear failed to diagnose a malignancy or indicate the possibility of a malignancy. False-positive: The cytologic smear was considered diagnostic for malignancy, suspicious for malignancy, or a differential diagnoses was given that included a malignant diagnoses; however, the corresponding surgical specimen did not reveal a malignant diagnosis.

All cases were reevaluated independently by two of the authors (A.M.A. and B.M.A.). Cases that had
both TP and CP were reevaluated separately, and the results were compared. Discrepant cases were reevaluated and categorized as sampling or interpretation errors. Clinical follow-up of patients without histologic correlation was obtained by reviewing the medical records.

RESULTS

Diagnosis

A total of 209 consecutive thyroid FNABs were examined during the study period. One hundred twenty-seven (61%) cases were prepared by TP and 82 (39%) by CP. Histologic correlation was available in 68 (33%) cases (32 TP and 36 CP) (Tables 1 and 2).

Cytologically, there were 19 cases with a malignant diagnosis, of which 18 cases were true-positive, and 1 case was false-positive. The false-positive case was diagnosed cytologically as a Hürthle cell neoplasm, but histologically the diagnosis of nodular hyperplasia with focal Hürthle cell changes was rendered.

Within the cytologically designated atypical/indeterminate category, there were 15 cases (6 TP and 9 CP) of follicular or Hürthle cell lesions. Thirteen cases were regarded as true-negative, as the histologic diagnosis was that of a benign lesion (7 cases of nodular hyperplasia and 6 cases of follicular adenoma). The remaining two cases were regarded as true-positive, as the corresponding histologic diagnosis was that of a follicular variant of papillary carcinoma.

There were 34 cases under the benign cytologic category. Both the cytologic diagnosis and the histologic examination were consistent with a benign diagnosis in 29 cases (24 cases of nodular hyperplasia or colloid nodule, 4 cases of follicular adenoma, and 1 case of lymphocytic thyroiditis). Five cases (three TP and two CP) were false-negative, and a cytologically benign, descriptive, nonspecific diagnosis was rendered; however, histologically a papillary carcinoma was identified in each case. On review, sampling error was responsible for all the false-negative diagnoses.

The overall sensitivity was 80%, whereas the sensitivity of CP versus TP was 87% and 70%, respectively. Meanwhile, the overall specificity was 98%, and the individual specificity for CP versus TP was 100% and 95%, respectively.

One hundred forty-one cases (95 TP and 46 CP) had no histologic correlation. For these cases, the cytologic diagnosis included 72 nodular goiter or cystic lesions (41 TP and 31 CP), 18 benign descriptive diagnoses (16 TC and 2 CP), 10 Hashimoto thyroiditis (7 TC and 3 CP), 6 follicular lesions (5 TC and 1 CP), and 35 nondiagnostic cases. Most (74%) of the nondiagnostic cases were prepared by TP (26 TP and 9 CP). Scant cellularity accounted for all the nondiagnostic cases (Table 3).
Sixty-eight (48%) of cases without histologic diagnosis had clinical follow-up, which ranged from 12–36 months. Evaluation of clinical outcome included periodic physical examination (55 cases), repeated fine-needle aspiration (10 cases), or ultrasound evaluation (3 cases). All patients followed a clinical course consistent with the corresponding cytologic or clinical diagnosis.

Twenty-four of the 36 cases that were prepared by CP and that had histologic correlation (Table 2) also had an accompanying TP slide. Of these TP, four were nondiagnostic due to low cellularity and seven cases based on the TP solely were diagnosed as follicular lesion. However, a more specific diagnosis of nodular hyperplasia was rendered on CP. There was no discrepancy in the diagnosis between TP and CP in the remaining 13 cases. Thus, the specificity and sensitivity of TP and CP were the same within this subgroup of cases.

### Cytologic Features

TP slide allowed assessment of the overall specimen cellularity but not individual passes of a FNAB as in CP. Specimen adequacy of thyroid FNABs prepared by CP and TP were assessed as previously discussed. The overall amount of colloid, and “watery” colloid specifically, was significantly diminished. The colloid tended to appear dense, markedly fragmented, or in droplets (Fig. 1). TP revealed more crowded, tight, tissue clusters and nuclear overlapping (Fig. 2A), compared with CP that showed flat, orderly sheets in a “honeycomb” arrangement (Fig. 2B). Loss of cellular preservation was especially obvious in the large aggregates in TP (Fig. 3). The peripheral edge of the preparation also commonly appeared blurred and poorly stained, an artifact associated with the TP technique. TP slides appeared to have more disruption of the cytoplasm, an increase in the number of naked nuclei, and a prominence of nucleoli. Nuclear grooves and “pseudoinclusions” were less apparent in papillary carcinoma (Fig. 4).

All the cytologic differences mentioned above were obvious, especially when comparing TP and CP from the same case. The major differences were the
Assessment of Cellularity
For specimen adequacy on CP, we require that there should be six or more clusters of cells, all of which are benign, on at least two slides before we accept a sample as adequate. Each cluster or group should contain at least 10 cells. Thus, when evaluating adequacy for TP specimens, we suggest a minimum of 12 clusters of follicular cells (each with at least 10 cells). Our study revealed that 74% of the unsatisfactory specimens were prepared by TP. This probably can be explained by the lack of on-site evaluation of specimen adequacy and by the lack of experience of the aspirator who obtained the specimen. In our institution, a sole TP with no smears commonly indicates that the aspirator was not the cytopathologist. Thus, although TP may reduce the number of slides with bloody background, we are still troubled by the nondiagnostic specimen. This issue will be resolved only through proper training and adequate cooperation with the cytology laboratory to allow on-site evaluation and/or specimen acquisition. Further, a nondiagnostic sample should be viewed within the Triple Test strategy. Thus, if the nodule is clinically or radiologically suspicious for malignancy, the nondiagnostic sample should be followed clinically, and additional sampling should be performed if deemed necessary. Additionally, nondiagnostic samples are commonly seen in cystic lesions. In such cases, the sample may consist...
solely of histiocytes and/or hemosiderin-laden macrophages. These cases should be evaluated cautiously. We render a descriptive nondiagnostic diagnosis and alert the clinician to the possibility of an underlying unsampled lesion. In our study, one true-positive case for papillary carcinoma was preceded by a nondiagnostic TP sample that consisted of only histiocytes. Additional follow-up sampling showed features that were consistent with a papillary carcinoma.

Of the overall 209 cases, there were 29 cases with a benign descriptive diagnosis. Of these, 83% (24 cases) were processed by TP compared with only 5 cases processed by CP. This difference between the two techniques may be explained by the lack of established criteria for evaluating TP-processed FNAB, minimal experience in evaluating artifacts associated with TP, and specimens of suboptimal adequacy.

**Colloid**

Assessment of the amount of colloid in the background plays an important role in the diagnosis of follicular lesions. Our study and that of previous authors showed that the amount of colloid on TP is diminished and, therefore, it is difficult to estimate its quantity. Colloid tended to appear dense, markedly fragmented, or in droplets. None of the cases in our study showed the diffuse watery colloid, which was probably lost during the TP processing. The approach that has been suggested by other authors, and which we apply to CP, is based on evaluation of two fundamental elements, the amount of colloid and the amount of follicular cells. Thus, such an approach may not be applicable when using TP. Generally, the cytologic evaluation of follicular lesions of the thyroid assumes the following: the more colloid, the more likely the lesion is benign; and the more cells, the more likely the lesion is neoplastic. Therefore, the lack of watery colloid in TP necessitates a modification in the approach to diagnosing a follicular lesion, and assigns a more critical role to the evaluation of the cytologic and architectural features of the cell clusters.

Sheets and clusters of benign follicular cells characterize colloid nodules processed by TP. In addition, histiocytes and rare “hard” fragments or droplets of colloid were identified. However, we noticed that as the sheets tended to get larger, they became crowded, tight, and difficult to evaluate. In contrast, follicular

**FIGURE 2.** Follicular cells aspirated from a multinodular hyperplasia prepared by the (A) ThinPrep Process and the (B) conventional preparation. Note more crowded tight tissue cluster and nuclear overlapping in (A) compared with (B) CP that showed flat, orderly sheets in a “honeycomb” arrangement (Papanicolaou stain, original magnification ×60).
neoplasms had numerous syncytial aggregates of follicular cells that formed microfollicles.

Architecture
Overall, TP revealed more crowded, three-dimensional, tight, tissue clusters with loss of cellular preservation within the center of the larger aggregates, increased disruption of the cytoplasm of the follicular cells, and more naked nuclei were encountered in the background. TP processing involves several steps that may contribute to this disruption in tissue architecture. These include the collection and suspension of the specimen in a fluid media, sample centrifugation, and dispersion of the cell suspension by rotation in a TransCyt filter assembly within the sample vial. Additionally, the suction effect of the pneumatic system to draw PreservCyt Solution through the filter membrane, the collection of the cells from the fluid media onto the filter membrane, and finally the transfer of the tissue onto a glass slide also may contribute to the occurrence of these artifacts.

Papillary Carcinoma
TP-processed samples showed alterations of some of the classic cytologic features of papillary carcinoma identified in CP. Well-formed papillary structures were rarely noted. Cell shrinkage artifact was a major problem encountered in TP in our study and in several other previous studies. Intranuclear cytoplasmic pseudoinclusions and nuclear grooves were less apparent.

The reported effect of the TP processing on the thyroid FNAB is limited. Biscotti et al. compared TP and CP of 41 thyroid FNABs obtained from surgically resected thyroid nodules. They experienced similar diagnostic accuracy with TP and CP. They reported that the TP slides exhibited less colloid that appeared as droplets. Additionally, they noted more cytoplasmic fragmentation, although associated with better nuclear detail.

Mesonero and Sickle reviewed 30 thyroid FNABs. Their study indicated that correlation between TP and CP diagnosis was seen in 90% of the cases. They also observed nuclear shrinkage and that colloid appeared markedly fragmented on TP slides. Meanwhile, Miller and colleagues in their study indicated that the TP-processing methodology gave diagnostic accuracy comparable to that of the conventional method. However, they noted that they had five major
discrepant diagnoses; in four, the CP was diagnosed as malignant, whereas the TP was benign, and the opposite was true in the remaining case.

Some studies have indicated that the TP technique may not be the ideal sole preparatory method for nongynecologic specimens, in particular for those lesions that contain fibroadipose or diagnostic stromal elements. Lee et al.22 prepared TP and CP from 100 FNABs obtained directly from surgical specimens removed from a wide variety of lesions. The authors noted that on TP, large tissue fragments were broken into pieces, epithelial-stromal relationships were more difficult to evaluate, single cells tended to aggregate into small clusters, shrinkage artifacts were more prominent, and nuclei appeared blander. Other studies also indicated that samples processed by TP are not consistently superior to CP.23,24

In summary, our study, although limited because of its small size, has shown that TP and CP of thyroid FNABs have nearly similar degrees of specificity and sensitivity. However, there are noticeable differences in the cytologic appearance of specimens processed by the two techniques. These morphologic differences must be recognized and may necessitate modification of diagnostic criteria established to evaluate CP when applied to TP. Further, the foundation for evaluating FNABs remains the on-site cytologic assessment by the pathologist, whenever possible. Using TP as the sole method of processing, and eliminating smears obtained at the time of on-site evaluation, may lead to an increased number of nondiagnostic samples in addition to potential diagnostic pitfalls. Awareness of the above-described findings and further studies that evaluate cytomorphologic features associated with TP processing are essential to avoid misleading diagnostic pitfalls in this era of medico-legal practice.

REFERENCES