The reliability of horizontally sectioned scalp biopsies in the diagnosis of chronic diffuse telogen hair loss in women

Rodney Sinclair,a Damian Jolley,b Rica Mallari,c and Jill Mageed
Melbourne, Victoria, Australia, and Manila, Philippines

Background: Chronic diffuse telogen hair loss is common in women. Paired 4-mm punch biopsy from the vertex scalp for horizontal and vertical sectioning is commonly used to distinguish between chronic telogen effluvium (CTE) and female pattern hair loss (FPHL). FPHL is now the favored term for androgenetic alopecia in women.

Objective and methods: To evaluate the reliability of a single horizontally sectioned scalp biopsy in the diagnosis of FPHL, 207 women presenting with chronic diffuse hair loss had three 4-mm punch biopsy specimens taken from immediately adjacent skin on the mid scalp, and all 3 biopsy specimens were sectioned horizontally. Findings were compared with 305 women who underwent two biopsies, with one sectioned horizontally and the other vertically. The terminal to vellus-like hair ratio (T:V) at the mid-isthmus level was used to diagnose FPHL (T:V ≥ 4:1), CTE (T:V < 8:1), or indeterminate hair loss (T:V = 5:1, 6:1, or 7:1). To correlate the histologic diagnosis with the clinical severity, a mid-scalp clinical grading scale was developed.

Results: Among the 305 women who had a single horizontal scalp biopsy, 181 (59%) were diagnosed as having FPHL, 54 (18%) having CTE, and 70 (23%) having indeterminate hair loss. Six hundred twenty-one horizontal biopsy specimens were assessed from 207 patients. On the basis of consensus over 3 biopsies, 159 (77%) were diagnosed as having FPHL, 44 (21%) having CTE, and the remaining 4 women (2%) as having indeterminate hair loss. Among these 207 women, 114 were assessed clinically as having stage 1 or 2 hair loss. Sixty-nine (60%) were diagnosed as having FPHL on the basis of triple biopsy, 42 (37%) having CTE, and 2 having indeterminate hair loss. Ninety-three were graded as having stage 3, 4, or 5 hair loss. FPHL was diagnosed in 90 women (97%), CTE in 2, and indeterminate hair loss in one. By using each single biopsy as the criterion for diagnosis, 398 (61%) were classified as FPHL, 99 (16%) as CTE, and 124 (20%) as indeterminate. In 493 biopsies (79%), the single biopsy conclusion was identical to the 3 biopsy conclusions. Where disagreement was seen (21%), most were classified as indeterminate, rather than as a wrong diagnosis (3.3%).

Conclusion: Application of these diagnostic criteria achieved accurate diagnostic definition in 98% of women with triple horizontal biopsies versus 79% with single horizontal biopsy. Ninety-seven percent of women with a mid-scalp clinical grade of 3, 4, or 5 were given a diagnosis of FPHL on triple biopsy. Scalp biopsy for diagnosis should be reserved for women with a mid-scalp clinical grade of 1 or 2. (J Am Acad Dermatol 2004;51:189-99.)

Women with androgenetic alopecia may present with a diffuse reduction in hair density on the crown, chronic increased telogen hair shedding, or both.1 Because the androgen-dependent nature has not been clearly determined in all cases of “androgenetic alopecia in women,” Olsen advocates use of the synonym “female pattern hair loss” (FPHL) preferentially. In
FPHL, the frontal hairline tends to be preserved, whereas men with androgenetic alopecia usually develop bitemporal recession early. Women who present with increased telogen hair shedding of more than 6 months' duration without any discernible reduction in hair density over the crown may have either chronic diffuse hair loss or FPHL and further investigation may be required to differentiate these two conditions.  

Chronic diffuse telogen hair loss may occur as a primary idiopathic condition (chronic telogen effluvium [CTE]) or be secondary to a range of occult nutritional, endocrine, autoimmune, or iatrogenic causes. Routine investigations usually identify secondary causes; however, primary CTE is a diagnosis of exclusion and a scalp biopsy is required to exclude early FPHL.  

The histologic hallmark of androgenetic alopecia in males and FPHL is miniaturization of hair follicles with a progressive transformation of terminal hair follicles into vellus-like follicles. Vellus-like follicles are defined as hairs with a hair shaft diameter of 0.03 mm or less and thinner than its inner root sheath. The shaft lacks pigment and a medullary cavity. Secondary vellus follicles are distinguished from primary vellus follicles by the presence of an arrector pili muscle and follicular streamers. Because this distinction is difficult on horizontal sections, all hairs are counted as vellus-like hairs whether primary or secondary to miniaturization from any cause.  

A 4-mm punch biopsy specimen taken from the mid scalp and sectioned horizontally is considered optimal for assessment of follicular miniaturization and detection of early FPHL. This site is selected for biopsy because FPHL is a patterned disorder that preferentially affects the mid scalp in women and spares the occipital scalp (Fig 1). Assessments of the number of vellus hairs is performed at the mid-isthmus level because that is an easily identified landmark where vellus hairs are most likely to be seen. Although vellus-like hairs may be more numerous higher in the papillary dermis, the possibility of artifact from trichostasis spinulosa is introduced (personal communication, D. A. Whiting, Dallas, Tex, June 2001).  

Based on studies in men, a reduction in the ratios of terminal to vellus-like hairs from greater than 6:1 to less than 4:1 is seen in androgenetic alopecia. Similar changes are seen in women with FPHL, and in a review of 219 horizontally sectioned biopsy specimens from women with FPHL, the average ratio of terminal to vellus-like hairs was 2.2:1.  

The cut points for diagnosis are a terminal/vellus (T:V) ratio of <4:1 for FPHL and >8:1 for CTE. However, it is not entirely clear how these cut points have been determined and validated.  

To determine the utility of horizontally sectioned scalp biopsy in the distinction between FPHL and CTE, a number of assumptions need to be tested. These include the following: that CTE is a distinct disease and not merely a precursor to FPHL; that pathologists can accurately identify and count vellus hairs in a horizontal scalp biopsy specimen; and that a single 4-mm scalp biopsy specimen is an adequate sample size. To test this last assumption, we estimated the reliability of a diagnosis for FPHL based on repeated observations from one woman.

**MATERIAL AND METHODS**

**Case selection**

Between 1997 and 2001, 545 females between the ages of 13 and 74 years were referred to a specialist hair clinic for investigation and treatment of hair loss of greater than 6 months' duration. All patients were examined by a specialist dermatologist (R. S.) to exclude other hair loss disorders, such as cicatricial alopecia, traction alopecia, trichotillomania, or alopecia areata. Screening hematologic investigations were performed to exclude thyroid disease, iron deficiency, zinc deficiency, systemic lupus erythematosus, and hormone dysregulation. All patients were offered a scalp biopsy as part of their assessment. Five hundred twelve patients agreed to scalp biopsy, and 33 patients declined the procedure.
Between 1997 and 1999, the protocol for scalp biopsy was to perform two 4-mm punch biopsies at the central mid scalp. One biopsy specimen was sectioned vertically and the other horizontally. The horizontally sectioned biopsy specimen was used to assess total hair count, vellus hair count, and terminal hair count. Three hundred thirty women agreed to paired scalp biopsy, and 25 patients declined the procedure.

Between April 2000 and February 2001, the protocol for scalp biopsy was to take 3 biopsy specimens from immediately adjacent skin on the mid scalp. All 3 biopsy specimens were sectioned horizontally, and the hair counts were performed in triplicate on each patient. Two hundred fifteen women agreed to triple scalp biopsy, and 8 patients declined the procedure. In view of the low sensitivity of the Ludwig scale to detect subtle changes in hair density, hair density was rated according to a modified scale that assessed the central part over the mid scalp (Fig 2).

Women with stage 1 and 2 hair loss tend to present with either episodic or continuous increased hair shedding. On direct questioning most agree there has been a reduction in the volume of the hair when held in a ponytail. Women with stage 3, 4, and 5 hair loss may present with either loss of volume of hair over the scalp with widening of the central part, episodic increased hair shedding, or continuous hair shedding.

An area on the central mid scalp was selected for biopsy and the skin marked with a surgical pen. If the patient parted her hair centrally, the site was shifted 1 to 2 cm laterally to avoid producing a scar in the part line. The skin was infiltrated with 2 mL of 2% lignocaine with 80,000 adrenaline (Astra Pharmaceuticals), and after a period of 10 minutes, a 4-mm disposable punch biopsy (Stiefel Laboratories, Inc, Coral Gables, Fla) was used to take the specimens. For the paired biopsy, 2 specimens are taken side by side and one submitted for routine vertical sectioning and the other for horizontal sectioning. For the triple-biopsy procedure, 3 plugs of immediately adjacent skin in an anterior-posterior line were taken and all were sent for horizontal sectioning (Fig 3). The wound is closed primarily with 2-0 nylon suture material (Ethicon), and the sutures are removed after 7 to 10 days.

**Histology**

The biopsy specimens were fixed in 10% neutral buffered formalin solution. For horizontal sectioning, the specimen is bisected horizontally 1 mm below the dermoeipidermal junction at the approximate entry of the sebaceous ducts into the follicle and the 2 portions embedded side by side, with the

---

**Fig 2.** Five-point visual analogue scale for assessment of female hair loss.
freshly cut sections face down in the block. Four slides containing 3 paired sections each were generally sufficient for evaluation of the horizontal section, but additional deeper sections were obtained if the initial section was too high or too low. Vertical sections were bisected vertically centrally. All sections were stained with hematoxylin and eosin for light microscopy (Fig 4). The total hair count was determined along with the total terminal hair count and the vellus-like hair count at the mid-isthmus level by a single observer (J. M.). The ratio of terminal to vellus hairs could then be calculated.

Mid-scalp clinical grading scale

Images for the mid-scalp clinical grading scale were chosen from a database of mid-scalp photographs developed to evaluate therapeutic response in women undergoing treatment for hair loss. The photographs were taken with the patient’s head in a stereotactic device and the camera mounted at a fixed distance from the scalp (Canfield Scientific, Inc., Fairfield, NJ). The 5-point visual analogue grading scale was tested for reproducibility. One hundred four patients used the scale to grade themselves before being seen in the clinic. Each patient was evaluated by a single investigator (R. S.) twice: on their first attendance and again between 2 and 4 weeks later. Mid-scalp photographs were taken between these two clinic visits using the stereotactic device, and this image was compared with the investigator’s assessment. The investigator was blinded as to the patient and photographic assessments. The scale was highly reproducible, with an identical assessment given by the investigator on both occasions in 90 patients (87%) and with the photographic image on 84 occasions (81%). Where discordance was seen, the deviation was only a single grade. Patients were less reliable using this tool for
self-assessment and tended to underestimate their hair loss.

**Statistical analysis**
Single biopsy results were compared with the overall result in several ways. In the first instance, the deviations of individual biopsy diagnoses from the diagnosis based on all 3 results were tabulated. Next, the intraclass correlation of the continuous measurement of the ratio of terminal to vellus hairs was computed, for all women combined and separately by diagnostic category. Finally, intraclass correlations were computed in a similar fashion for alternative diagnostic criteria, such as the number of vellus hairs and the number of terminal hairs.

**RESULTS**

**Paired biopsy**
Among the 305 women who had a single horizontal scalp biopsy, 181 (59%) were diagnosed with RPHL (T:V <4:1), 54 (18%) with CTE (T:V >8:1), and 70 (23%) as indeterminate (T:V = 5:1, 6:1, or 7:1). Baseline clinical grading of the severity of the hair loss was not available for these women.

**Triple biopsy—diagnostic agreement indices**
Among the 207 women, the hair density was assessed clinically as stage 1 in 57 women (27.5%), stage 2 in 57 (27.5%), stage 3 in 66 (31.9%), stage 4 in 15 (7.2%), and stage 5 in 12 (5.7%). Histologically 159 (77%) were diagnosed with FPHL on the basis of consensus over 3 biopsies (one or more biopsies with a T:V <4:1), 44 (21%) as having CTE (one or more biopsies with a T:V >8:1 and no biopsy with T:V <4:1), and the remaining 4 women (2%) were classified as “indeterminate” diagnosis (no biopsy with a T:V <4:1 or T:V >8:1).

Among the 159 women with FPHL, 32 (20%) were assessed clinically as having stage 1 hair density, 37 (25%) as stage 2, 64 (40%) as stage 3, 14 (9%) as stage 4, and 12 (7.5%) as stage 5. Among the 44 women diagnosed histologically with CTE, 23 (52%) were assessed as stage 1, 19 (43%) as stage 2, 2 (5%) as stage 3, and none were assessed as either stage 4 or 5. Among the 4 women with indeterminate hair loss, 2 were assessed clinically as stage 1, 1 as stage 2, and 1 as stage 4.

Using each single biopsy as the criterion for diagnosis, there were 621 biopsies done on the 207 women. Of these 621 biopsies, 398 specimens (61%) would be classified as FPHL on the basis of the T:V ratio alone, 99 (16%) would be classified as CTE, and the remainder (124; 20%) would have been classified as indeterminate (Table I).

The most obvious departure from agreement is the classification of “indeterminate”; this was seen only for 4 women (12 biopsies) on the basis of 3 biopsies, whereas a total of 124 biopsies were classified as “indeterminate” on the basis of a single
biopsy. The errors favored the CTE diagnosis proportionally more than FPHL. This is consistent with the greater relative proportion of FPHL to CTE in the entire population.

Overall, in 493 biopsies (79%), the single-biopsy conclusion was identical to the triple-biopsy conclusion. Of the remaining 21% of biopsies for which a disagreement was seen, most were classified as indeterminate, rather than a wrong diagnosis. Of 477 “true” FPHL biopsy results, only 3.3% were wrongly classified as CTE; in the other direction, of 132 “true” CTE biopsy results, none was judged as FPHL on the basis of a single biopsy. All the remaining errors were “true” FPHL or “true” CTE that failed to show a definitive diagnosis from the single biopsy.

It was not possible to calculate realistic estimates of “sensitivity” or “specificity” because these measures must be based on an independent “gold standard” diagnosis. In the current design, the result of the triple-biopsy “gold standard” was based on the combined results of the 3 biopsies and is therefore not independent of any one of the constituent results. In an attempt to resolve this problem, “diagnosis” based on the other two biopsies was computed so that a more independent comparison could be made (Table II). These results, taken from two other biopsies independently assessed, are more realistic “gold standards” than the result based on all 3 biopsies. These calculations were performed for all 207 patients in the database and comparisons were made for the 621 single-biopsy diagnoses. The results are presented in Table III.

The most striking feature of these results is the similarity in the number of “indeterminate” results—in both cases 20% of all biopsies, although fewer than half of them agree on the indeterminacy—particularly when this is compared with the triple-biopsy “gold standard” diagnosis shown in Table I. This probably reflects the avoidance of indeterminate results for the “gold standard” and the fact that the “other two” diagnosis has been made solely on the basis of the ratio of the total T to total V.

Unsurprisingly, disagreement is far more prevalent in Table III than in Table I. It is not surprising because the “gold standard” in Table I used infor-

---

### Table I. Diagnosis on the basis of one biopsy versus three biopsies

<table>
<thead>
<tr>
<th>Diagnosis based on single biopsy</th>
<th>Diagnosis based on three biopsies</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPHL</td>
<td>CTE</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>FPHL</td>
<td>397</td>
</tr>
<tr>
<td>CTE</td>
<td>16</td>
</tr>
<tr>
<td>Indet</td>
<td>64</td>
</tr>
<tr>
<td>Total</td>
<td>477 (75%)</td>
</tr>
</tbody>
</table>

CTE, Chronic telogen effluvium; FPHL, female pattern hair loss; Indet, indeterminate.

### Table II. Variables used to compute diagnosis on the basis of other two biopsies

<table>
<thead>
<tr>
<th>Biopsy</th>
<th>Total</th>
<th>Terminal</th>
<th>Vellus</th>
<th>m_Total</th>
<th>m_Terminal</th>
<th>m_Vellus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>44</td>
<td>32</td>
<td>12</td>
<td>90</td>
<td>74</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>44</td>
<td>38</td>
<td>6</td>
<td>90</td>
<td>68</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>36</td>
<td>10</td>
<td>88</td>
<td>70</td>
<td>18</td>
</tr>
</tbody>
</table>

### Table III. Diagnosis based on one biopsy versus that of the other two biopsies

<table>
<thead>
<tr>
<th>Diagnosis based on single biopsy</th>
<th>Diagnosis based on other two biopsies</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPHL</td>
<td>CTE</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>FPHL</td>
<td>361</td>
</tr>
<tr>
<td>CTE</td>
<td>8</td>
</tr>
<tr>
<td>Indet</td>
<td>42</td>
</tr>
<tr>
<td>Total</td>
<td>411 (100%)</td>
</tr>
</tbody>
</table>

CTE, Chronic telogen effluvium; FPHL, female pattern hair loss; Indet, indeterminate.
information from the single biopsy and thus can be expected to be more likely to agree. In Table III, the overall agreement rate was 466/621 (75.0%). Taking the “other two” as the preferred diagnosis, the correct proportion (“sensitivity”) for FPHL was 88% and for CTE was only 62%. Once again, most of the incorrect results were classified as indeterminate, and the respective misdiagnosis rates were 1.9% for FPHL and 3.6% for CTE.

One measure of agreement between observers is the $k$ statistic, which measures the chance-adjusted agreement between two diagnoses. From Table III, the expected agreement, based on chance and using the marginal totals, is 48.5%; the observed agreement (75%) is clearly better than this, but the value of $k$ computed from these is 0.50. It would be misleading to place a confidence interval (CI) around this estimate, since the data remain clustered groups of 3 results for each patient. Nevertheless, we can interpret the value: “good” agreement is generally accepted for $k > 0.75$ and “poor” agreement for $k < 0.4$. In the present case, the agreement can be interpreted as at best moderate.

### Intraclass correlation

Regardless of the cut points used for diagnosis, we can measure the consistency between repeated quantitative measurements on the same subject using standard statistical methods. The “intraclass correlation coefficient” (ICC) is one measure of the reliability of a single quantitative measurement. By using data from all patients, we can use a one-way analysis of variance (ANOVA) method to estimate the ICC for the ratio T:V for the FPHL data. These results are shown in Table IV.

There were 614 biopsy results included in the analysis. In 7 cases, the vellus hair count was zero, causing the ratio measure to be undefined. This happened for more than one biopsy for only one patient (Oct #10), causing her results to be effectively dropped from the study. From Table IV, the ICC for the T:V ratio was 0.53, with a 95% CI from 0.45 to 0.60. This shows only moderate agreement for the quantitative measure. The within-person standard deviation (SD) was 3.3, which is very large given that the diagnostic cut point for FPHL is the T:V ratio ≤ 4.0.

Note that the estimated reliability of a mean T:V ratio based on 3 biopsies is 0.77. This is much smaller and less subject to error than results based on one biopsy only. The ICC result is much less when results are computed within diagnostic category. In both cases the ICC is less than the value for all subjects combined.

For FPHL patients, the ICC is 0.30, which is relatively poor agreement. The within-patient SD has been reduced to 1.66, which is more acceptable than for all patients combined; the lower value is due to the “noisier” CTE patients being removed.

For CTE patients, results are very poor indeed. The ICC is 0.16, indicating very little agreement at all, and the within-patient SD is a huge 6.6. The results of the ANOVA show a $P$ value of .04, which indicates only marginally significant association within patients’ results. In other words, the ratio measures of T:V are almost consistent with random assignment of patient identification numbers among the CTE patients. The poor agreement for CTE patients is not surprising, considering the nature of the ratio measure being used. Ratios are always unstable when their denominator (ie, the count of vellus hairs, in this case) is low or close to zero. Small
changes in the V value can induce huge variation in the criterion ratio measure. The relative instability of the ratio measure is manifested by the distributional shape of the measure across the range of patients, as can be seen from Fig 3.

In Fig 3 the individual quantiles of the ratio are shown as circles, and they have been sorted from lowest (zero) to highest (~40) and plotted against the corresponding quantile of a normal distribution with similar mean and SD. The straight line shown is the shape of the theoretical symmetric normal distribution. The “skewedness” of the distribution of T:V ratios is evident from the curvature in the sequence of circles. The extremes to the high values are much higher than predicted by a normal shape, and the lower end is “squashed” into a region above zero.

Alternative diagnostic criteria

Other possible predictors of diagnostic outcome may be considered, other than the ratio criterion. For example, the number of vellus hairs by themselves may differentiate between FPHL and CTE conditions. This does not appear to be the case if one examines the joint relation between the numbers of terminal and vellus hairs (Figs 5 and 6). On these charts, the terminal hair count from each biopsy specimen has been plotted (x-axis) against the matching vellus hair count (y-axis). In both plots, straight lines are used to identify the criterion regions of FPHL (the upper region), CTE (the lowest region), and indeterminate (the middle region). The “gold standard” diagnosis is shown for each biopsy by using the plotting symbols of cross (FPB) and triangle (CTE). In Fig 7, log scales have been used on both axes to display more clearly the “action” in the lower left corner. In none of the figures is there evidence of any maximum absolute number of vellus hairs that could distinguish between FPHL and CTE diagnoses. There is nowhere a horizontal line could divide without misclassification of the crosses and triangles.

Clinicohistologic correlation

Paired biopsy. Baseline clinical grading of the severity of the hair loss was not available for the 305 women who underwent paired biopsy with a single horizontal scalp biopsy.

Triple biopsy—diagnostic agreement indices. Among the 207 women, hair density was assessed clinically as stage 1 in 57 women (27.5%), stage 2 in 57 (27.5%), stage 3 in 66 (31.9%), stage 4 in 15 (7.2%), and stage 5 in 12 (5.7%). Among the 159 women with FPHL, hair density was assessed clinically as stage 1 in 32 patients (20%), stage 2 in 37 (25%), stage 3 in 64 (40%), stage 4 in 14 (9%), and stage 5 in 12 (7.5%). Among the 44 women diagnosed histologically with CTE, 23 (52%) were assessed as stage 1, 19 (43%) as stage 2, 2 (5%) as stage 3, and no women were assessed as either stage 4 or 5.
the 4 women with indeterminate hair loss, 2 were assessed clinically as stage 1, 1 as stage 2, and 1 as stage 4.

Of the 57 women assessed clinically with stage 1 hair loss, 23 (56%) had FPHL diagnosed histologically, 23 (40%) had CTE, and 2 were indeterminate.

Of the 66 women with stage 3 hair loss, 64 (97%) had FPHL and 2 had CTE. Of the 15 women with stage 4
hair loss, 14 (93%) had FPHL and one was indeterminate. Of the 12 women with stage 5 hair loss, all had FPHL histologically.

**DISCUSSION**

There is no “gold standard” for the diagnosis of FPHL. A clinical diagnosis based on the Ludwig classification may underestimate early FPHL and overestimate advanced FPHL. A scalp biopsy is useful to identify women with early FPHL who present with increased scalp hair shedding, but little or no reduction in hair volume over the crown (stages 1 and 2); scalp biopsy can also help to distinguish between CTE and FPHL. Biopsy is also useful in identifying women with fibrosing alopecia in a pattern distribution that can mimic advanced FPHL. It has been suggested that some women who present with a marked reduction of hair volume over the crown and who have evidence of inflammation and hair follicle fibrosis on histology have patterned lichen planopilaris rather than FPHL with fibrosis.8

For a horizontally sectioned scalp biopsy specimen to become established as a “gold standard” for biopsy, several characteristics need to be determined. These include the ideal site for biopsy, the size of the biopsy specimen and the number of specimens required, and the histologic criteria for a diagnosis of FPHL. In addition, it must be determined whether specialist dermatopathologists can accurately identify vellus hairs and count them.

This article estimates the reliability of a diagnosis for FPHL in women based on repeated observations from the same woman and whether the T:V ratio is a more reliable measure than the absolute number of vellus hairs.

The number of cases in which a conclusive diagnosis was unable to be made (indeterminate) was reduced from 23% with a single horizontal biopsy to 2% with triple biopsy. Because it is relatively simple to persuade a patient to have three 4-mm punch biopsies from the same site rather than 2, and relatively difficult to explain to a patient that the biopsy he or she has had was noninformative and must be repeated, we advocate triple biopsy. The need for anesthetic, the time taken to do the procedure, and the morbidity associated with the procedure are unaltered.

FPHL in women is a diffuse process. Material differences in the total hair count, the vellus hair count, and the degree of fibrosis were seen in skin immediately adjacent to the vertex scalp. The degree of disagreement was sufficient to shift the biopsy into a different diagnostic class in 21% of patients. In most of the cases where disagreement was seen, one of the biopsy specimens was classified as indeterminate, whereas the other 2 showed either CTE or FPHL. However, a wrong diagnosis was seen in 3.3% overall, where one biopsy specimen if read in isolation would have led to a diagnosis of CTE, but one or both of the remaining biopsy specimens from immediately adjacent skin showed FPHL. This was seen in approximately 15% of all the cases diagnosed as CTE.

We found that the ratio measure for differential diagnosis between FPHL and CTE appears to be the most appropriate to use, but it suffers from very poor reproducibility. The SD of a single within-person ratio measure is 1.6 for FPHL patients and as high as 6.6 for CTE patients. Such large SDs are unacceptable when a cut point between 4 and 8 is used to distinguish the diagnoses.

This article also correlates the histologic diagnosis with a novel validated 5-point mid-scalp clinical grading scale. Ninety-seven percent of women graded as 3, 4, or 5 had FPHL on biopsy, and we believe CTE can be excluded in these women on a clinical basis alone. Whether a biopsy is still required to exclude patterned lichen planopilaris is a matter for conjecture; however, those women often have associated symptoms, such as scalp pain, scalp itch, perifollicular erythema, or hyperkeratosis.

Sixty percent of women with clinical mid-scalp grades of 1 or 2 had FPHL on biopsy. It is not possible in our experience to exclude FPHL without a scalp biopsy among women who present with the classic symptoms and signs of CTE, such as diffusely increased hair shedding (including the occipital scalp), reduction in volume of ponytail thickness, bitemporal recession, and absence of widening of the central part. Although women with hair shedding of less than 6 months’ duration were not included in this study, many women diagnosed with acute telogen effluvium with no obvious trigger can be shown to have FPHL on biopsy. In the early stages of FPHL, it is not uncommon for women to have one or more episodes of increased hair shedding that terminates spontaneously. Many of these women also describe an apparent recovery between episodes.

If a biopsy is indicated to diagnose or exclude FPHL, to take account of the diffuse nature of the process we advocate that a triple-biopsy procedure be performed on the vertex scalp and all specimens sectioned horizontally because application of established diagnostic criteria achieves accurate diagnostic definition in 98% of women with triple horizontal biopsies versus 79% with single horizontal biopsy.

**REFERENCES**