Agreement in histologic assessments of the pituitary pars intermedia in aged horses

Dianne McFarlane, DVM, PhD; Lisa M. Miller, DVM, PhD; Linden E. Craig, DVM, PhD; Noel O. Dybdal, DVM, PhD; Perry L. Habecker, VMD; Margaret A. Miller, DVM, PhD; Jon S. Patterson, DVM, PhD; Alastair E. Cribb, DVM, PhD

Objective—To evaluate concordance among veterinary pathologists in the assessment of histologic findings in the pars intermedia of pituitary gland sections from aged horses with mild signs suggestive of pituitary pars intermedia dysfunction (PPID).

Sample Population—10 pituitary glands from aged horses.

Procedure—7 pathologists were provided with signature, clinical signs, and a single H&E-stained pituitary gland section from 10 aged horses with mild signs suggestive of PPID. Pathologists described histologic findings for each section and stated whether findings were consistent with PPID. Agreement among pathologists and with antemortem diagnostic test results was calculated.

Results—Overall, only fair agreement was found among the pathologists as to which horses had histologic findings consistent with disease (mean ± SE kappa value, 0.34 ± 0.069). Interpretation of individual sections varied, with minimal agreement (4 or 5/7 pathologists) for 5 of 10 sections evaluated. Postmortem assessment was in agreement with an antemortem endocrine diagnostic test result 79% of the time.

Conclusions and Clinical Relevance—Validation of antemortem diagnostic testing for PPID in horses often relies on testing hypothalamic-pituitary-adrenal axis responsiveness or measurement of endogenous plasma concentrations of proopiomelanocortin-derived peptides, such as ACTH.

The 19-hour (overnight) dexamethasone suppression test (DST) is considered the standard method of antemortem PPID diagnosis. In clinically normal horses, IM administration of dexamethasone decreases release of ACTH from the pars distalis, resulting in a serum cortisol concentration of < 1 µg/dL (27.59 nmol/L) 19 hours after dexamethasone administration. Horses with PPID fail to suppress serum cortisol concentrations as a result of ACTH production from the pars intermedia. Originally, this test was reported to have a sensitivity and specificity of 100%. However, a recent report suggests that the reliability of the test has been overestimated. In addition, seasonal variation in response to dexamethasone has been documented. Clinically normal ponies and horses had normal overnight DST results in spring, but when the same animals were tested in the fall, they failed to suppress. Although it has not been critically assessed, a loss of feedback inhibition by glucocorticoids may be a late event in the disease progression and the high sensitivity originally reported may reflect a case selection bias toward horses with advanced disease. Other limitations of the overnight DST include the need for multiple-day sample collection and the potential exacerbation of laminitis in horses with current or historical laminitis.

As a result of the limitations of the overnight DST, other antemortem tests such as endogenous ACTH or α-melanocyte-stimulating hormone (α-MSH) plasma concentration measurements, thyrotropin-releasing hormone response test, or combined hormone response tests have been the subject of recent, more critical evaluation. Emphasis has been placed on their potential use for early diagnosis. Sensitivity and specificity of the diagnostic tests have been established by comparing results to clinical signs and postmortem diagnosis. In aged horses with advanced disease, the presence of clinical signs including hirsutism, laminitis, muscle atrophy, abnormal fat distribution, polydip-
sia and polyuria, hyperhidrosis, and secondary infections may be sufficient for diagnosis. In horses with early disease, however, clinical signs are often difficult to distinguish from changes observed with normal aging. Therefore, heavy emphasis is placed on postmortem diagnosis of PPID for establishing accuracy of ante-mortem diagnostic testing protocols. The purpose of the study reported here was to evaluate the concordance among pathologists in the histologic assessment of pars intermedia pathologic changes in pituitary gland sections from aged horses.

Materials and Methods
Animaless—All samples were collected in accordance with the guidelines of the Canadian Council on Animal Care. Ten aged horses and ponies with mild clinical signs consistent with PPID were included in the study. Following clinical evaluation and testing, horses were euthanized with pentobarbital (125 mg/kg, IV).

Endocrinologic tests—Nine of the 10 horses received an overnight DST prior to euthanasia. One of the tested horses had received triamcinolone acetate for treatment of airway disease several weeks prior to admission, and results of the overnight DST were not useful. A second horse was clinically in distress and was euthanized prior to receiving an overnight DST.

Prior to euthanasia, plasma α-MSH concentrations were measured in the 10 horses by use of a radioimmunoassay previously validated for use in horses. Plasma concentrations of > 91 pmol/L were previously reported to be diagnostic of PPID. This reference value was established with 63 horses (38 unaffected and 25 affected horses) by use of overnight DST results as the standard. Sensitivity and specificity were reported to be 88% and 84.2%, respectively. However, seasonal variations in plasma α-MSH concentration and overnight DST results were not considered when determining these reference range values.

In a study on a group of 11 healthy horses and 14 healthy ponies monitored periodically over 1 year, the mean ± SD plasma concentration of α-MSH in the spring, summer, and winter was much lower (10.9 ± 3.6 pmol/L) than reported for control horses by Horowitz et al (59.6 ± 91.6 pmol/L), whereas plasma concentrations in the fall were similar in the 2 studies (49.2 ± 49.3 pmol/L vs 59.6 ± 91.6 pmol/L). By use of mean plasma concentrations ± 2 SDs as cutoff values, plasma α-MSH concentrations of > 19 pmol/L in the spring, summer, or winter and plasma concentrations of > 148 pmol/L in the fall were considered diagnostic of PPID. These seasonally specific reference range values were used in this study.

Histologic slide preparation and evaluation—Seven consecutive, 5-μm-thick, formalin-fixed, paraffin-embedded sagittal sections of pituitary glands from 10 aged horses were mounted on glass slides treated with 3-amino-propyltriethoxysilane diluted in acetone and stained with H&E. Seven diplomates of the American College of Veterinary Pathology were recruited for participation in the study. Pathologists had a stated interest in endocrine disease, were at an institute where endocrine research in horses is actively conducted, or both.

Each pathologist received a set of slides consisting of a single section from pituitary glands of the 10 horses or ponies and a description of the signalment and clinical signs (ie, muscle atrophy [n = 6], persistent guard hairs [3], heaves [2], degenerative joint disease [2], history of laminitis [2], weight loss [2], obesity with cresty neck [1], lethargy and possible narcolepsy [1], metritis [1], periorbital...
squamous cell carcinoma with conjunctivitis [1], and tachycardia and jugular pulses [1]). Pathologists were instructed to describe the histologic findings for each section. In addition, pathologists were asked to give an opinion as to whether histologic findings of each section were diagnostic for PPID. Pathologists were unaware of the results of endocrine diagnostic tests at the time of histologic evaluation.

Statistical analysis—Results of histologic evaluation were analyzed for agreement among pathologists by establishing multi-evaluator kappa values with the use of a software program. Agreement of histopathologic diagnosis with antemortem diagnostics (either plasma α-MSH concentration or overnight DST result) was calculated. Values of $P < 0.05$ were considered significant.

Results

Horses and ponies in this study had a mean age of 23.8 years (range, 15 to 32 years). The group consisted of 4 mares and 6 geldings of various breeds (ie, 3 Quarter Horses, 1 pony crossbred, 1 pony–Quarter Horse crossbred, 1 Thoroughbred, 1 Thoroughbred crossbred, 1 Standardbred, 1 Standardbred crossbred, and 1 Appaloosa). Of the 10 horses or ponies, the most common clinical signs suggestive of PPID were muscle atrophy or weight loss ($n = 7$), persistent guard hairs ($2$), and laminitis ($2$). One of the laminitic animals was a grossly obese pony; the other laminitic horse had chronic metritis. None of the horses or ponies had hirsutism.

Table 1—Summary of the pathologists’ (A through G) histologic descriptions of sections of pituitary glands from horses or ponies ($n = 10$) with mild clinical signs of pituitary pars intermedia dysfunction (PPID).

<table>
<thead>
<tr>
<th>Pathologist</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>NORM</td>
<td>NORM</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>HYT</td>
<td>NORM</td>
<td>HYPL</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>HYPL</td>
<td>HYPL</td>
<td>HYPL</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>HYPL</td>
<td>NORM</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>HYPL</td>
<td>HYPL</td>
<td>HYPL</td>
<td>HYPL</td>
<td>HYPL</td>
<td>HYT</td>
<td>HYPL</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>ADMT*</td>
<td>PPID*</td>
<td>ADMT*</td>
<td>NORM</td>
<td>PPID*</td>
<td>ADMT*</td>
<td>NORM</td>
<td>PPID*</td>
<td>NORM</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>PPID*</td>
<td>HYPL</td>
<td>ADHD*</td>
<td>HYPL</td>
<td>HYPL</td>
<td>PPID*</td>
<td>HYPL</td>
<td>PPID*</td>
<td>HYPL</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>NORM</td>
<td>PPID*</td>
<td>NORM</td>
<td>HYPL</td>
<td>NORM</td>
<td>NORM</td>
<td>NORM</td>
<td>NORM</td>
<td>NORM</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>HYT</td>
<td>HYT</td>
<td>HYPL</td>
<td>ADMT*</td>
<td>ADMT*</td>
<td>HYPL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*For analysis of findings in this study, histologic findings considered positive for PPID. HYPL = Pars intermedia hyperplasia. NORM = Histologic findings consistent with those of clinically normal aged horses or ponies. ADMT = Pars intermedia adenomatous hyperplasia (consistent with PPID). HYT = Pars intermedia melanotrophic hypertrophy. AMD = Pars intermedia adenoma (consistent with PPID).

Table 2—Comparison of antemortem diagnostic test results and postmortem diagnosis in horses or ponies ($n = 10$) with mild clinical signs of PPID.

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antemortem Month of testing</td>
<td>Dec</td>
<td>Jan</td>
<td>May</td>
<td>Jun</td>
<td>Jul</td>
<td>Feb</td>
<td>May</td>
<td>Sep</td>
<td>Jul</td>
<td>Jan</td>
</tr>
<tr>
<td>Plasma $\alpha$-MSH (pmol/L)</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Neg</td>
</tr>
<tr>
<td>Overnight DST</td>
<td>ND</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Neg</td>
<td>NP</td>
</tr>
<tr>
<td>PPID diagnosis</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
<td>Pos</td>
<td>Neg</td>
</tr>
<tr>
<td>Postmortem PPID*</td>
<td></td>
<td>4 (57.1)</td>
<td>7 (100)</td>
<td>6 (85.7)</td>
<td>5 (71.4)</td>
<td>2 (28.6)</td>
<td>2 (28.6)</td>
<td>4 (57.1)</td>
<td>4 (57.1)</td>
<td>7 (100)</td>
</tr>
<tr>
<td>Agreement (%)</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>6 (85.7)</td>
<td>5 (71.4)</td>
<td>4 (57.1)</td>
<td>4 (57.1)</td>
<td>7 (100)</td>
</tr>
</tbody>
</table>

*Number of evaluators out of 7 with histopathologic findings of PPID. Number of evaluators out of 7 with histologic findings in agreement with antemortem results. 
$\alpha$-MSH = $\alpha$-Melanocyte–stimulating hormone. Neg = Negative. Pos = Positive. DST = Dexamethasone suppression test. ND = Not diagnostic. NP = Not performed.
All 7 pathologists provided written reports for each histologic section examined (Figure 1). Five pathologists commented that either the irregular plane of a section or the lack of multiple sections for each animal limited them. Five pathologists made the comment that PPID was a clinical diagnosis only. Four limited their evaluation of the sections to histologic descriptions, providing no final diagnosis. Therefore, a finding of pars intermedia adenoma, adenomatous hyperplasia, nodular hyperplasia, or microadenomatous pars intermedia was considered consistent with PPID for the purpose of analysis in this study (Table 1).

Kappa values were calculated to assess agreement among evaluators. Perfect agreement among evaluators would have a kappa value equal to 1; no agreement beyond what would be expected to occur by chance alone would have a kappa value of 0. Kappa values can be interpreted as very good (> 0.8), good (0.61 to 0.8), moderate (0.41 to 0.6), fair (0.21 to 0.4), and poor (< 0.2) agreement. The mean ± SE kappa value for agreement among evaluators for interpretation of all 10 pituitary gland sections was 0.34 ± 0.069, indicating that agreement was only fair but significantly (P < 0.001) greater than that expected as a result of chance alone. Interpretation of individual sections varied, with minimal agreement (4/7 or 5/7 pathologists) in 5 of 10 sections evaluated.

On the basis of seasonally specific cutoff values, 3 horses were considered to have PPID. On the basis of overnight DST results, 1 additional horse was considered to have PPID. Therefore, 4 horses had an ante-mortem diagnosis of PPID by either overnight DST results or an increased α-MSH plasma concentration. Overall, histologic assessment was in agreement with antemortem diagnostic test results 79% of the time (Table 2).

Discussion

Diagnosis of PPID in horses is not straightforward. Pituitary pars intermedia dysfunction is a progressive, degenerative condition; therefore, identification of a discrete starting point for the disease is likely to be difficult. Early disease recognition is obscured by the overlap between phenotypic changes accompanying age and disease. Histologic changes in pituitary glands of healthy aged horses or ponies are not well described and likely overlap with lesions observed in early disease. In a recent study of 100 healthy horses, pathologic changes of pituitary glands (cysts, hyperplasia, microadenomas, or adenomas) and size (weight, length, width, and height) were reported to increase with age. Changes in the appearance of the pars intermedia when physiologically activated, such as seasonally in the fall, have not been investigated. Evidence of pathologic changes that accompany disease progression in individual horses or ponies is currently lacking.

In reading the comments provided by the pathologists in our study, it was clear that no agreement existed on the histopathologic definition of PPID. One pathologist clearly stated that diagnosis was reserved for horses with gross enlargement of the pars intermedia resulting in compression of the hypothalamus and pars distalis or pars nervosa, or both, such that functionality of these regions would be compromised. These horses would be expected to have overt clinical signs, such as hirsutism. Others considered the disease to include adenomas that did not compress adjacent lobes, whereas some included diffuse hyperplasia and focal microadenomas. Clinical signs may be expected to be subtler in these affected animals and likely the result of an increase in plasma concentration of pro-opiomelanocortin-derived peptides. A concise definition of the histologic findings consistent with PPID is needed to facilitate accurate communication between pathologists and clinicians.

In our study, overall agreement among pathologists in the histologic interpretation of the 10 pituitary gland sections was fair (kappa value, 0.34). However, for 3 of the 10 sections, only 4 of 7 pathologists agreed, and for 2 sections, only 3 of 7 pathologists agreed, indicating considerable variation. This may reflect the difficulty in differentiating early histopathologic changes from normal changes associated with physiologic stimulation. In histologic evaluation of tissue from 2 horses considered positive for PPID on the basis of plasma α-MSH concentration, 100% agreement was found among the pathologists. Also, strong agreement was found among pathologists in histologic evaluation of tissue from 3 horses that did not have endocrine test results indicative of PPID (agreement among 6, 7, and 6/7 pathologists). Of these 3 horses, 2 horses were among those with the lowest plasma α-MSH concentration. The least agreement was found among pathologists in histologic evaluation of tissue from horses that had intermediate plasma α-MSH concentrations. The lack of agreement among pathologists in assessment of tissues from horses with intermediate antemortem endocrine test results indicates that postmortem evaluation is not likely reliable in diagnosing early PPID.

It may be that diagnosis of PPID should not be made without antemortem endocrine testing and overt phenotypic changes consistent with disease. If this is true, the use of postmortem histologic evaluation as the standard for evaluation of antemortem diagnostic tests is inappropriate. Identification of early disease status will need to be addressed in a novel manner, as the currently available diagnostic tests are not adequately validated for use in early disease. Accurate diagnosis of early disease may not prove possible as a result of the progressive nature of PPID.

Results of our study highlight the need for further definition of the histologic changes that occur in pituitary glands of horses with normal aging, with seasonal physiologic stimulation, and with PPID. In addition, until the histologic changes that accompany early disease are clarified, antemortem diagnostic tests validated by use of postmortem histologic evaluation as the standard should be viewed cautiously. It is likely that with currently available diagnostic testing methods, early disease may go undetected or be incorrectly diagnosed. Therefore, careful consideration of clinical signs and progression of those signs as well as repeated endocrine function testing would be prudent in aged horses in which early PPID is suspected.


d. SAS, version 8.2, SAS Institute Inc, Cary, NC.

References


