Identification of Vertical Growth Phase in Malignant Melanoma

A Study of Interobserver Agreement

Nuala C. McDermott, MRCPath,1 Donal P. Hayes, DipRCPath,2 Mohammed H. Al-Sader, DipRCPath,1 John M. Hogan, FRCPat.h,2 Caitriona Barry Walsh, FRCPat.h,1 Elaine W. Kay, MD,1 and Mary B. Leader, MD1

Key Words: Malignant melanoma; Radial growth phase; Vertical growth phase; Breslow thickness; Clark level; Observer variation; Agreement

Abstract

Ninety-four H&E-stained slides of malignant melanoma were circulated to 6 pathologists in 2 university departments. For each slide, the growth phase of the lesion, Breslow thickness, and Clark level were determined by each observer. The aims of the study were to evaluate agreement between nonspecialist pathologists in identifying the vertical growth phase in malignant melanoma and to compare agreement for the growth phase with agreement for Breslow thickness and the Clark level. Our results show that although overall agreement for the growth phase is moderate, agreement between experienced observers is good. In fact agreement for the growth phase among this group was equal to the agreement for Breslow thickness and the Clark level. Our results show that although overall agreement for the growth phase is moderate, agreement between experienced observers is good. In fact agreement for the growth phase among this group was equal to the agreement for Breslow thickness. Overall agreement for Breslow thickness also was good but for the Clark level was only fair. These findings suggest that if the predictive value of the vertical growth phase proves to be robust, it will be used with an acceptable level of accuracy in routine diagnostic practice.

One of the features of the increase in cutaneous malignant melanoma in recent years has been a rise in the proportion of thin lesions being diagnosed. These lesions are not infrequently a source of considerable diagnostic difficulty as the distinction between a benign melanocytic proliferation with atypical features and a thin melanoma can be problematic. Several recent studies have shown that interobserver agreement for diagnosis can be quite variable when a spectrum of melanocytic lesions is assessed,1-4 but concordance tends to improve after the adoption of standardized criteria,1,3 As the quest for diagnostic precision continues, so does the discussion and refinement of histologic criteria for distinguishing the different categories of melanocytic lesions. Meanwhile, in the context of patient management, it is important that there is good agreement between practicing pathologists in distinguishing the lesions that are capable of metastasis from those that pose no further threat to the patient once they have been removed.

In 1984, Clark et al5 applied the concept of stepwise tumor progression to malignant melanomas. They proposed that melanocytic tumors do not progress uniformly in space and time but evolve in a stepwise fashion, each step being characterized by the acquisition of characteristics not manifest in the preceding lesion.5 These steps can be related to the histologic categories used in the diagnosis of melanocytic lesions. Clinical and histologic features characteristic of each step were described by Clark et al5 and form the basis of the criteria of Elder and Murphy6 and of the Cancer Research Campaign melanoma panel (CRC panel).1

According to Clark et al,5 an established malignant melanoma evolves through a period of radial growth from which the vertical growth phase arises. The radial growth...
Materials and Methods

Ninety-four H&E-stained slides of malignant melanoma were circulated to 6 pathologists in 2 university departments; 3 were experienced pathologists (minimum 10 years post MRCPath), and 3 were less experienced (pre-MRCPath). For each slide, the growth phase of the lesion, Breslow thickness, and Clark level were determined by each observer. In an attempt to simulate as closely as possible normal diagnostic practice, pathologists were requested to use their usual method of assessment for all variables. To allow for the possibility that an individual pathologist might not be as familiar with the concept of growth phase as with Breslow thickness and Clark level, each observer was given a recent article relating to growth phase and the problems associated with its application. The criteria for differentiation between radial and vertical growth phases are clearly outlined therein. Invasive melanoma in the radial growth phase is characterized by the presence of melanoma cells singly or in small groups (usually <10 cells wide) within the papillary dermis. These cells are cytologically similar to the melanocytes in the overlying epidermis. Mitoses are rare. A prominent lymphocytic infiltrate often is associated with this dermal component. The vertical growth phase is characterized by the impression of an expansile nodule within the dermis. This nodule should be a minimum of 10 cells wide. A nodule 25 cells wide is a definite indication of vertical growth phase. The component cells usually show a greater degree of cytologic atypia than do the melanocytes in the overlying epidermis. Mitoses may be seen, and the nodule seems to dominate other groups of melanoma cells within the dermis. A dermal nodule of intermediate size (10–25 cells wide) may represent a vertical growth phase component if it seems to dominate other dermal nodules or nests, its component cells are cytologically more atypical, and mitoses are present.

We calculated \( \kappa \) statistics using the statistical package Stata (version 5) (Stata, College Station, Tex). This method of analysis provides an indication of the strength of agreement between observers. It requires that data be in categorical form. Clark level and growth phase data are already in this form, but the data for Breslow thickness was continuous. To apply \( \kappa \) statistics, each measurement of Breslow thickness was allocated to 1 of the following categories: less than 0.76, 0.76 to 1.5, or more than 1.5 mm Table II. The following is a guideline to the interpretation of the \( \kappa \) value: for \( \kappa < 0.20 \), the strength of agreement is poor; for 0.21 to 0.40, fair; for 0.41 to 0.60, moderate; for 0.61 to 0.80, good; and for 0.81 to 1.00, very good.
Values of $\kappa$ for Agreement

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Experienced</th>
<th>Inexperienced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth phase</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breslow thickness (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.56</td>
<td>0.68</td>
<td>0.40</td>
</tr>
<tr>
<td>&lt;0.76</td>
<td>0.71</td>
<td>0.68</td>
<td>0.76</td>
</tr>
<tr>
<td>0.76-1.5</td>
<td>0.73</td>
<td>0.69</td>
<td>0.76</td>
</tr>
<tr>
<td>&gt;1.5</td>
<td>0.66</td>
<td>0.64</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>0.71</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>Clark level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.34</td>
<td>0.51</td>
<td>-</td>
</tr>
<tr>
<td>Level 1</td>
<td>0.25</td>
<td>0.14</td>
<td>0.24</td>
</tr>
<tr>
<td>Level 2</td>
<td>0.49</td>
<td>0.64</td>
<td>0.37</td>
</tr>
<tr>
<td>Level 3</td>
<td>0.14</td>
<td>0.38</td>
<td>-0.19</td>
</tr>
<tr>
<td>Level 4</td>
<td>0.38</td>
<td>0.59</td>
<td>0.17</td>
</tr>
<tr>
<td>Level 5</td>
<td>-0.0021</td>
<td>-0.0041</td>
<td>-</td>
</tr>
</tbody>
</table>

To present the data in a way that allows a more detailed visual examination of agreement between pairs of observers, a scatter plot matrix was assembled for each variable [Figure 11, Figure 21, and Figure 31]. In such a matrix, total agreement will result in all observations lying on a line with a slope that is 45°. The amount of scatter about this line indicates the extent of disagreement. Variation in the slope of the line is an indication of bias (patterned disagreement).

**Results**

**Growth Phase**

The results for the growth phase show that overall agreement was moderate ($\kappa = 0.56$). However, when considering experienced observers separately, the level of agreement between them is good ($\kappa = 0.68$), although the level of agreement between the inexperienced raters is only fair ($\kappa = 0.40$).

**Breslow Thickness (Categorized Data)**

The overall level of agreement was good ($\kappa = 0.71$). The best overall agreement was seen for lesions of more than 1.5 mm ($\kappa = 0.77$) and the worst in lesions 0.76 to 1.5 mm ($\kappa = 0.66$). The level of agreement for Breslow thickness was lower between experienced observers and higher between inexperienced observers ($0.68$ vs $0.76$). Both values correspond to a good level of agreement. The scatter plot matrix for Breslow thickness shows a small number of outlying points that represent extreme disagreements. These points are largely the result of 4 cases that were a constant source of disagreement for this variable.

**Clark Level**

For Clark level, overall agreement was fair ($\kappa = 0.34$). The best result was seen in level 2 in which agreement was moderate ($\kappa = 0.49$) and the worst in level 3 in which agreement was poor ($\kappa = 0.14$). However, if the results for Clark level are evaluated in terms of experienced and inexperienced observers, for all categories, the agreement was higher in the experienced groups ($\kappa = 0.51$), and for Clark level 2, agreement among experienced pathologists was good ($\kappa = 0.64$). Inexperienced observers have a much lower level of agreement for all categories except Clark level 1.

**Discussion**

The importance of accurate identification of melanoma in the vertical growth phase is emphasized in the CRC report: “If further studies confirm that the absence of a vertical growth phase is synonymous with recognising an inability to metastasise, the most important prognostic assessment would become not the assessment of thickness but recognition of vertical growth phase.” The significance of the vertical growth phase is that the evidence to date suggests that it represents a distinct step in the biologic evolution of a melanoma; the clinical and histologic features by which it is recognizable are the morphologic expression of change at a fundamental level. This is illustrated by studies that have demonstrated that unlike radial growth phase lesions, the cells of vertical growth phase melanoma are readily established in tissue culture, do not require exogenous growth factors for proliferation, and readily form tumors in nude mice.7

If the radial growth phase is synonymous with inability to metastasize, then assessment of prognostic variables is valuable only in the vertical growth phase lesions. It has been shown that when the vertical growth groups are evaluated separately, thickness loses some of its predictive ability (radial growth phase lesions are predominantly thin). In these circumstances, other histologic features, such as mitotic rate, tumor infiltrating lymphocytes, and the presence of regression, along with clinical variables, such as lesional site and patient sex, assume predictive value. The recognition of the
vertical growth phase, therefore, alters the criteria that we use to assess prognosis. It also is logical that it should be considered when evaluating other markers of potential prognostic significance, such as proliferative activity (Ki67, MIB1) or tumor progression (nm23, CD44, or members of the integrin family), in future studies. In particular, it underlines the desirability of using true survival as a determinant of outcome when assessing such markers. While the relationship of a marker with other features of prognostic significance such as thickness is undoubtedly interesting, the true significance of expression of any potential prognostic marker can be determined only by its ability to predict survival.

We conducted the present study to determine whether there was concordance between pathologists in applying the concept of radial and vertical growth phase. An article referring to the work of Clark et al acknowledged that the study was large and well executed but stated that “the theory does need validation in other centers to ensure that a ‘vertical’ growth phase pattern is recognized uniformly by many pathologists.” The importance of a high level of agreement is emphasized by De Wit et al and also in the report of the CRC panel. Our results show good agreement between experienced observers. It is particularly notable that among this group, agreement for growth phase equals agreement for Breslow depth, which is widely regarded as the most reliable and objective of the routinely used prognostic indicators. It is interesting but not surprising that while the less experienced observers had high levels of agreement for the objective measurement of thickness, concordance for subjective variables was higher among the more experienced group.

The report of the CRC panel also recognizes that the level of agreement achieved by specialists may not be representative of general histopathologic practice. The high level of agreement for growth phase among experienced nonspecialist observers in the present study augurs well for its use as an indicator of prognosis in routine diagnostic practice. It also is recognized that agreement tends to improve as criteria are more precisely defined. The criteria outlined by Elder and Murphy and the CRC panel are clear and precise and should contribute substantially to further improving agreement for this variable.

**Conclusion**

This study of agreement between nonspecialist pathologists in identifying the vertical growth phase in malignant melanoma demonstrated good concordance between experienced nonspecialist pathologists. These findings suggest that if the criteria prove to be robust, they are likely to be used accurately and may contribute substantially to our ability to predict the outcome of malignant melanoma.
From the 1Department of Pathology, Royal College of Surgeons in Ireland, St Stephen’s Green, Dublin, Ireland, and 2Department of Pathology, Cork University Hospital, Cork, Ireland.

Address reprint requests to Dr McDermott: Department of Pathology, Royal College of Surgeons in Ireland, St Stephen’s Green, Dublin 2, Ireland.

Acknowledgments: We acknowledge the assistance of Professor Alan Kelly, PhD, Department of Statistics, Dublin University, and Mr Ronan Conroy, Department of Epidemiology, Royal College of Surgeons in Ireland.

References


