

Measurement techniques for melanoma: a statistical comparison

C J Calder, A P Campbell, S R Plastow

Abstract

Inter- and intra-observer variation in measuring the depth of invasion of malignant melanomas was assessed using three different techniques: eye-piece graticule, stage Vernier, and projection image analysis. Significant variation was found for all methods but was least pronounced with the stage Vernier. It is recommended that this should be the preferred technique for routine use.

It has been clearly established that the depth of invasion of the dermis is the single most important prognostic factor in determining the outcome of any melanoma.¹ It is therefore essential to have reliable, accurate techniques for the measurement of this. This study examines intra-observer variation in assessment of the Breslow thickness² using three different measurement techniques: the eye-piece graticule, the stage Vernier, and projection image analysis, and inter-observer variation using the same three techniques performed by two pathologists (CJC and APC).

Methods

Consecutive sections of 75 invasive malignant melanomas were taken from the file. The identification was obscured and each slide was assigned a number between 1 and 75. The sections had all been cut in the same laboratory at about 5 μ m and stained with haematoxylin and eosin. The sections were examined jointly by the pathologists to determine the deepest point of the lesion, which was then marked. This was done to ensure that the only variables were the method of measurement used and the pathologist doing the measuring.

The melanomas were then measured by each of the three techniques:

1—An eye-piece graticule on an Olympus H2 microscope; this had been calibrated using a standard slide.

2—The Vernier scale on the stage of the same microscope.

3—Projection image analysis—the melanoma was projected onto a digitising tablet (Graph-ec, UK) coupled to an Apple microcomputer. The Breslow thickness was measured by touching the stylus to the upper and lower points of the projected image allowing the linear distance to be calculated by a standard computer program. The full microscope field was projected to about 15 cm in diameter and the measurements were calibrated by using a standard scale placed on the microscope stage (a plastic ruler). A mean of four readings was taken.

Having measured the 75 melanomas using each of the three techniques, the results were analysed using standard statistical methods. Pairwise comparisons were made, using paired Student's *t* tests to detect any differences between measurement techniques and differences between the two observers using the same technique. Analysis using a two way ANOVA was also performed³; this has the advantage of detecting any interaction between a particular method and the observer. It was felt that any method with a pronounced interaction between observer and technique would be more prone to inaccuracies and hence would not be the technique of choice for routine use in the reporting of malignant melanoma.

The statistical analyses were repeated excluding those melanomas greater than 2 mm in thickness (as these all fall into a poor prognostic group) to see if there was a greater inherent variation in measuring smaller distances.

Results

The two way ANOVA showed that there was no significant interaction between the individual pathologist and the techniques used. A slight but non-significant difference in the techniques was seen using this analysis ($p > 0.1$).

The results from the paired *t* tests are summarised in the table. Significant variation was observed with one of the pathologists using the eye-piece graticule and the image analysis techniques. This variation was not present when the melanomas over 2 mm thick

Department of
Histopathology,
Wycombe General
Hospital,
Buckinghamshire
C J Calder
A P Campbell

Photobiology Unit,
Institute of
Dermatology, UMDS,
London
S R Plastow

Correspondence to:
Dr CJ Calder,
Department of Pathology,
The Medical School,
Birmingham B15 2TJ.

Accepted for publication
11 June 1990

Approximate *p* values from paired *t* tests

Comparison tested:		<i>p</i> Value (all melanomas)	<i>p</i> Value (melanomas < 2 mm)
Graticule <i>v</i> Vernier	(Pathologist 1)	0.33	0.8
Vernier <i>v</i> projection	(Pathologist 1)	0.15	0.43
Graticule <i>v</i> projection	(Pathologist 1)	0.15	0.41
Graticule <i>v</i> Vernier	(Pathologist 2)	0.15	0.30
Vernier <i>v</i> projection	(Pathologist 2)	0.075	0.20
Graticule <i>v</i> projection	(Pathologist 2)	0.025*	0.13
Pathologist 1 <i>v</i> pathologist 2	(graticule)	0.025*	0.001*
Pathologist 1 <i>v</i> pathologist 2	(Vernier)	0.125	0.011*
Pathologist 1 <i>v</i> pathologist 2	(projection)	0.003*	0.001*

*Significant variation

were excluded from the data. No other significant intra-observer variation was seen.

There was significant interobserver variation when using the eye-piece graticule and the image analysis system. No significant difference between pathologists was seen when using the stage Vernier. When measuring the thinner melanoma sections (<2 mm), significant variations were seen with all three techniques. The stage Vernier technique, however, showed the least significant interobserver variation.

Discussion

In this study we have tried to assess which of three possible techniques gives the most reproducible measurements of depth of dermal invasion and which would therefore be the preferred technique for routine use. Variables were restricted to the technique used and the pathologist doing the measuring by reviewing the melanomas before starting the study. Despite doing this, significant interobserver variation was seen with all techniques. This implies the presence of an inherent subjective element in each method. The reason why variation was seen least with the Vernier method is not clear. It may be that for

each edge of the melanoma the microscope is adjusted and then the reading is taken, rather than attempting to assimilate all the information at once as is done with the eye-piece graticule. Problems when using the image analyser are that pathologists are more used to looking at a microscopic image than a projected image and that measurements are automatically logged into the computer, making fine adjustments difficult.

It seems, therefore, that for all depths of malignant melanoma the stage Vernier will give measurements which are more likely to be reproducible among pathologists and therefore among centres than either of the other methods and as such is the preferred technique for routine use.

We thank Dr MJ Turner, Department of Histopathology, Wycombe General Hospital, and Dr KA Fleming, Nuffield Department of Pathology and Bacteriology, John Radcliffe Hospital, Oxford, for their helpful advice.

- 1 Breslow A, Cascinelli N, Van Der Esch EP, Morabito A. Stage I melanoma of the limbs: assessment of prognosis by levels of invasion and maximum thickness. *Tumori* 1978;64:273-84.
- 2 Breslow A. Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. *Ann Surg* 1970;172:902-8.
- 3 Daniel WW. Biostatistics: a foundation for analysis in the health sciences. 4th edition. New York: John Wiley and Sons, 1987:273-323.