OBJECTIVE: To test the interobserver and intraobserver reproducibility of the standard terminology for description and diagnosis of melanocytic lesions in in vivo confocal microscopy.

DESIGN: A dedicated Web platform was developed to train the participants and to allow independent distant evaluations of confocal images via the Internet.

SETTING: Department of Dermatology, University of Modena and Reggio Emilia, Modena, Italy.

PARTICIPANTS: The study population was composed of 15 melanomas, 30 nevi, and 5 Spitz/Reed nevi. Six expert centers were invited to participate at the study. Intervention Evaluation of 36 features in 345 confocal microscopic images from melanocytic lesions.

MAIN OUTCOME MEASURE: Interobserved and intraobserved agreement, by calculating the Cohen kappa statistics measure for each descriptor.

RESULTS: High overall levels of reproducibility were shown for most of the evaluated features. In both the training and test sets there was a parallel trend of decreasing kappa values as deeper anatomic skin levels were evaluated. All of the features, except 1, used for melanoma diagnosis, including roundish pagetoid cells, nonedged papillae, atypical cells in basal layer, cerebriform clusters, and nucleated cells infiltrating dermal papillae, showed high overall levels of reproducibility. However, less-than-ideal reproducibility was obtained for some descriptors, such as grainy appearance of the epidermis, junctional thickening, mild atypia in basal layer, plump bright cells, small bright cells, and reticulated fibers in the dermis. Conclusion The standard consensus confocal terminology useful for the evaluation of melanocytic lesions was reproducibly recognized by independent observers.