Comparison of the Use of Standardized Diagnostic Criteria and Intuitive Clinical Diagnosis in the Diagnosis of Common Viral Warts (Verrucae Vulgaris)

Richard Young, MBBS; Damien Jolley, MSc (Stats), MSc (Epidemiol); Robin Marks, MBBS, MPH

Objective: To assess the diagnostic value of standardized diagnostic criteria compared with the clinical intuitive diagnosis for verrucae vulgaris.

Design: A blind comparison with cross-over of experienced dermatologists using either intuitive clinical diagnosis or else recording the presence of standardized diagnostic criteria in verrucae vulgaris.

Setting: Clinical outpatient department in a major teaching hospital.

Patients: Fifteen patients aged 6 to 60 years with 21 verrucae vulgaris and 24 control lesions.

Intervention: Dermatologists examined the lesions and recorded either the diagnosis by the usual clinical intuitive method or else whether standardized diagnostic criteria were present in the lesion of each patient. Each dermatologist did 1 only of the 2 methods for each lesion with the other doing the alternative method for the same lesion, neither of them having previously been aware of any of the diagnoses.

Main Outcome Measure: The sensitivity, specificity, and positive predictive, negative predictive, and relative values for each diagnostic approach were compared.

Results: Use of intuitive clinical diagnosis resulted in 100% specificity, sensitivity, and positive predictive, negative predictive, and relative values for both the verrucae vulgaris and the control lesions. Recording the presence of the diagnostic criteria reduced to a varying extent all the parameters of diagnostic accuracy with site—fingers and hands, elbows, and knees—being the most efficacious of the criteria.

Conclusions: These data indicate that care is required in relying too much on standardized criteria as the basis for clinical diagnosis, even for a lesion as simple as a wart. They also reinforce the value of clinical experience obtained from repeated exposure to multiple variants of a disease in the traditional bedside teaching model.

Arch Dermatol. 1998;134:1586-1589

From the University of Melbourne, Department of Medicine (Dermatology), St Vincent’s Hospital (Melbourne), Fitzroy, Victoria, Australia.
SUBJECTS AND METHODS

STUDY SUBJECTS

Subjects were seen in the Dermatology Outpatient Department at St Vincent's Hospital, Melbourne, Australia. They were selected from among hospital staff members and their family, and from patients of the Department of Dermatology. All of the cases were patients with warts who were being seen or treated within the clinic setting, or who had had them diagnosed previously.

CRITERIA

Two more criteria that did not achieve the same statistical significance were believed by clinicians to be clinically important and were added to the list of 4 diagnostic criteria previously tested. They comprised a papillomatous surface to the lesion and black dots just beneath the surface of the lesion, representing punctate hemorrhages. Two experienced dermatologists were used in this study so it was anticipated there would be little, if any, misunderstanding of the meaning of such terms as papillomatous or any of the others used in the diagnostic criteria.

EXAMINATION

All of the lesions were examined first by the senior physician (R.M.) who knew the medical history and treatment and who confirmed and recorded the diagnosis, circled each lesion, and numbered it with a felt-tipped pen. Each lesion was then examined by the 2 experienced dermatologists.

The first dermatologist examined the lesion and recorded which of the criteria from the checklist were present or absent. This dermatologist was asked not to record a diagnosis. The second dermatologist then examined the same lesion without conferring with the first dermatologist. This time a diagnosis was made for the lesion using findings from the clinical examination, without referring to the presence or absence of the diagnostic criteria (intuitive diagnosis). The examiners then moved onto the next lesion but reversed their roles, so that the dermatologist who had given the clinical diagnosis now recorded the presence of diagnostic criteria and the other, a clinical diagnosis. This continued until all lesions had been seen by both dermatologists alternately recording diagnostic criteria only or an intuitive diagnosis only.

No patient was seen twice by any one dermatologist. The aim was not to test interobserver variation, but to compare the sensitivity and specificity for the diagnosis of warts with the use of the diagnostic criteria with that of using intuitive clinical diagnosis.

ANALYSIS

All the data were entered onto a relational database (FileMaker Pro for Macintosh, Claris Corporation, Santa Clara, Calif). The original diagnoses by the senior clinician (R.M.) were converted to binary form, ie, a wart or not a wart. Then each of the intuitive diagnoses and presence of the diagnostic criteria for each lesion were compared with the original diagnosis. The sensitivity, specificity, and positive predictive, negative predictive, and relative values were then calculated using the version 6.1 SPSS software package (SPSS Inc, Chicago, Ill). The relative value provides a guide to the overall discriminatory usefulness of individual criterion. It is derived by adding the sensitivity and the specificity and subtracting 100. It has a maximum value of 100.

RESULTS

PATIENTS AND LESIONS

Fifteen patients aged 6 to 60 years were seen. A total of 45 lesions were examined, 21 of which were common viral warts. There were 24 control lesions that were selected from other lesions already present on the skin of the subjects with warts. These 24 lesions comprised 7 seborrhoeic keratoses, 5 melanocytic nevi, 4 solar keratoses, 3 scars, 2 traumatic lesions, 2 raised psoriatic lesions, and 1 milial cyst. Viral warts in all stages of evolution were selected, including several that were partially treated.

COMPARISON OF INTUITIVE DIAGNOSIS WITH THE LIST OF DIAGNOSTIC CRITERIA

The dermatologists had a sensitivity and specificity of 100% compared with the known diagnosis when using their usual clinical or intuitive method of diagnosis of the viral warts (Table). Under these conditions, the clinicians also correctly diagnosed 100% of the control lesions.

On the other hand, recording the presence of diagnostic criteria alone and not offering a diagnosis, reduced the sensitivity, specificity, and positive predictive, negative predictive, and relative values to varying degrees in almost all circumstances. The only instance where there was a 100% sensitivity and a 100% negative predictive value was with the presence or absence of a “discrete margin” in the lesions being examined. Although this criterion managed to pick up all the warts present, it was not specific to these lesions and, thus, the specificity for discrete margin was only 25%.

The only one of the standardized diagnostic criteria to approach a relatively high degree of compatibility with intuitive clinical diagnosis was site—fingers and hands, knees, and, elbows. This criterion had the stron-
The aim of these studies originally was to create a checklist of diagnostic criteria that could be used as a valid instrument in a large epidemiological survey, possibly even increasing the accuracy of diagnosis. As it turned out, each of the criteria in the final list proved to be less discerning than the clinicians who were involved in the study and merely used their clinical or intuitive diagnosis. We had assessed various combinations of these criteria in the previous studies with the combination of 2 (flesh colored and hyperkeratotic) giving the highest relative value. Increasing the combination of criteria used in diagnosis increased the specificity but substantially decreased the sensitivity and thus the relative value overall. Thus, we have still avoided specifying what combination of diagnostic criteria had to be fulfilled to qualify as a “case” of viral wart.

This study raises a number of questions regarding both the diagnosis of common viral warts and, on a more general level, the derivation of diagnostic criteria in trying to simplify or validate the clinical diagnostic process. From the results obtained, it seems that experienced dermatologists have absolutely no difficulty in diagnosing common warts, as one would expect. How they achieve this has not been specifically determined, but it is assumed that it is the combination of rigorous training, a perceptive eye, and years of experience. Use of diagnostic criteria, on the other hand, led to a less satisfactory outcome. It was suggested at the outset of the study that it is a simple matter to diagnose a common wart clinically, and that a checklist would not improve diagnostic accuracy. This has been borne out.

The question of why the use of the checklist of criteria was not as good as rapid intuitive diagnosis must be raised. Were the selected criteria wrong or were there not enough? In an attempt to avoid this possibility, the definition of a common wart was derived from a major international textbook of dermatology used by a large number of dermatologists throughout the world. This was combined with criteria suggested by 43 practicing dermatologists and dermatology registrars in Victoria, Australia. By doing so, it was believed that the list would be comprehensive and also based on clinical practice.

Was the wrong short list selected from the large number of criteria derived initially? The previous analysis and results published showed that there were 4 useful diagnostic criteria for diagnosing common warts. No limit was set in the analysis on how many criteria could be included. They were derived using multivariate logistic regression analysis that then excluded criteria that were not statistically significant.

The method of deriving each diagnostic criterion’s statistical strength was by comparison with the criterion standard diagnosis by the study dermatologist. The same dermatologist was used to select warts, and other lesions, in both studies. This would tend to overcome the possibility that there was a diagnostic problem in the initial selection process for the diagnostic criteria that created problems for the second study. This study was loosely based on 2 others: the American Rheumatism Society study on systemic lupus erythematosus, and the UK Working Party study on atopic dermatitis in which similar methods were used. Both of these groups managed to devise what has been reported to be a workable list of criteria that could be used with confidence within statistical parameters.

Visual discrimination is a complex neurophysiological and psychological process. This study suggests that there is probably vastly more to the process than we realize, certainly more than we are able to reduce to a small number of simplistic salient features. Even for a lesion as simple as a viral wart, it appears that we either see features that have not been able to be easily recorded in our clinical descriptions or else that we build on simple salient features and create neurologically more complex patterns for recognition that are not easily described by the diagnostic criteria. This is analogous to the neural networks computer programs that attempt to replicate some of the human neurophysiological information processing.

There is no doubt that clinical diagnostic criteria are important to standardize diagnosis and produce data from different centers that are comparable. Nevertheless, this study highlights the care that is required. Although it is praiseworthy to attempt to standardize classification of disease, it is important not to underestimate the process of visual discrimination, even for a lesion as simple as a wart.

Finally, the traditional training in dermatology is based on years of exposure to vast numbers of examples of common cutaneous conditions in an environment that is conducive to informed learning. This is the standard approach to learning clinical skills. There are increasing moves in medical education toward less time in the clinical environment and more time in laboratories with computer-generated interactive learning programs. These are absolutely dependent on formal diagnostic criteria in the development of the programs. Thus, our study also stands as warning that there may be more basis to the traditional training in learning der-
matology and other branches of clinical medicine than we realize. Careful thought is required before we rely too much on the modern reductionist approach to learning clinical skills and rely too little on the lessons learned from history on the value of the bedside.

Accepted for publication September 22, 1998.

We would like to thank dermatologists Andrew Michaelides, MBBS, and Chris Baker, MBBS, for being the independent observers for this project. We would also like to thank the patients and their families who gave us their time to assist us with our study.

Corresponding author: Robin Marks, MBBS, MPH, Department of Medicine (Dermatology), St Vincent’s Hospital (Melbourne), Fitzroy, Victoria 3065, Australia.

REFERENCES