

Evaluation of clinical diagnostic tests in dermatology

Marina JOVANOVIĆ*

Clinic of Dermatovenereology Diseases, Clinical Center of Vojvodina, Novi Sad, Serbia

*Correspondence: Marina JOVANOVIĆ, E-mail: serbjdermatol@nadlanu.com

UDC 616.5-07



Dermatology is primarily a visual discipline (1,2). For example, a group of students can easily be recognized holding their hands behind their backs when examining a patient from a distance. When faced with dermatological conditions they are not familiar with, non-specialists prefer a tool that helps them to search for images of skin conditions by body region or morphology, rather than by the condition's name. Using the Map of Dermatology, non-specialists can achieve a diagnosis or differential diagnosis by browsing the resultant image sets and compare them to the clinical presentation at hand (2). Teledermatology techniques can be reliably applied even to retrospective diagnosis (3,4). However, although the vast majority of dermatologic conditions are accessible for examination by visual inspection, dermatologists also touch and sometimes smell the skin as well as look at it (1). The fundamental point that palpation in itself has diagnostic utility in dermatologic diagnosis, even when the visual stimulus is absent, has recently been confirmed. This may explain some of the reservations regarding teledermatology (1). Often being the sole diagnostic tool in patients care, clinical examination has a crucial role in dermatology. Moreover, without clinicopathologic correlation, the histopathologic diagnosis would extremely be limited (5).

Issues

So, what do we do now? A patient has a disease, but experiences an illness. We need to determine how often the medical history and physical examination provide a highly accurate diagnosis. We also have to determine whether specific historical or physical examination findings exist, and whether they can predict the disease (6). If a disease is characterized by more than one basic lesion, only consideration of other criteria will permit us to consider the right diagnosis. For example, when patients with atopic dermatitis present with plaques, papules, vesicles, pustules,

excoriations, or various combinations of these lesions, pruritus, although nonspecific, will probably be the best diagnostic symptom of atopic dermatitis (4). Thus, better and more comprehensive maps should be developed, based on differential diagnosis algorithms/trees, taking into consideration many other aspects of patients and their lesions, in addition to images (2,7). The Decision Tree is one among different approaches to create a classification model as an optimal predictive model that will further be a referential diagnostic tool for physicians. Data mining is an important part of this information technology (7).

Classification of clinical signs and symptoms

Medical diagnosis may be considered as a categorization task. At least 2 different processes, by which this categorization task may be accomplished, are operative. In "analytic processing" the clinician reaches a diagnosis by identifying and combining clinical signs and symptoms. The diagnosis is made after careful evaluation of all clinical parameters exhibited in a particular case. This strategy has traditionally been strongly proposed and considered as an expert diagnostic hallmark. Instructions aimed at advancing the "analytic processing" can improve the accuracy of typical cases. On the other side, "nonanalytic processing" makes use of the similarity between previously encountered examples and the present case. As clinicians gain experience, they rely less on "analytic processing" and more on "similarity-based processing". Thus, analytic approach to diagnosis is replaced by a more holistic or similarity-based approach, as one gains significant clinical experience. If reliance on prior examples is related to the expertise level of the participants, one might find a greater reliance on rules or less reliance on similarity to prior cases with less expert participants or beginners. Instructions that foster "analytic processing" would have little effect on

the performance of inexperienced participants, since they are already inclined toward this strategy. However, instructions fostering a similarity-based approach, would have positive effects on their performance. Thus, inexperienced participants who are encouraged to use a similarity-based approach would exhibit a higher rate of accuracy for similar cases, regardless of typicality (8).

Assessment of clinical examination

Clinical examination can and should be assessed as a diagnostic test within the discipline of clinical epidemiology. The main strategy of this discipline is improving the patient-clinician interaction (6). It is important in dermatology, especially regarding precision, accuracy and utility of alternative diagnostic strategies, such as epiluminescence and digital imaging, as well as differences in practice patterns between dermatologists and nondermatologists.

All diagnostic tests, including clinical examinations in dermatology, need high quality evidence about the level of precision and accuracy. Precision (reliability, reproducibility, repeatability) refers to agreement. When related to clinical examination, it is called "observer agreement". Intraobserver agreement refers to the agreement the same observer has at 2 different evaluations of the same patient. "Interobserver agreement" refers to the agreement between 2 different observers, who independently (blinded), within minutes or at most hours, examine the same patient. If the same diagnosis is obtained, then it is precise, if not, the diagnosis is imprecise. "Accuracy" answers whether the diagnosis is correct or incorrect, by correlating or not with the truth. Imprecise diagnoses (at least one) will inherently result in inaccuracy, though inaccuracy cannot be predicted from precision estimates alone (6).

We cannot always know the truth regarding the status of the disease. A "gold standard" or reference standard test should be chosen to represent the truth. It will be the best test available. A good example of a gold standard is biopsy specimen for histologic evaluation to diagnose skin cancer. The accuracy of the test (diagnosis) is measured with its sensitivity (a proportion of patients that will have a positive test among patients who actually have the disease), specificity (a proportion of patients that will not have a positive test among patients who do not have the disease), positive predictive value (the proportion of patients with positive test results, who are correctly diagnosed) and neg-

ative predictive value (the proportion of patients who do not have the disease, if they do not have positive test results). The positive predictive value will increase, while the negative will decrease as the prevalence of disease in the study population increases. Likelihood ratio is derived not from the prevalence, but from sensitivity and specificity, and expresses the odds that the diagnostic test (diagnosis) has in a patient with the disease, as opposed to a patient without the disease (6).

If examinations performed by different examiners are imprecise, then the diagnosis is examiner-dependent. The patient's treatment depends on who performs the examination. On the other side, inaccuracy will result in misdiagnosis and affect management and treatment outcome. Precision, clinical agreement, namely consensus, may be necessary more than application of a gold standard, if a gold standard does not exist, or the diagnostic test (clinical examination) and the gold standard are the same. For example, if dermatologists gain consensus, the results of their highly precise examinations can be used to categorize the disease. Ideally, data are categorized as present or absent, then chance-corrected measures of agreement can be used, e.g., k -statistics (k values the range from -1 representing complete disagreement, 0, agreement by chance alone, and 1 - complete agreement). Clinical follow-up is another strategy that can overcome the difficulty in applying the gold standard. Clinical follow-up is the main strategy that should be used in patients with negative results of examinations. Reexaminations should be done in reasonably short periods. Moreover, it is worth knowing that the small number of incorrectly categorized misdiagnoses can dramatically affect sensitivity rates (6).

Diagnostic skills in dermatology

There is a continuing trend of non-dermatologists to treat skin diseases. A recent study assessed the accuracy of clinical diagnoses made by physicians, not specifically trained in dermatology, and by dermatologists using histopathological diagnosis as the "gold standard" (5). The results clearly demonstrated that dermatologists diagnosed twice as many cases correctly, compared with non-dermatologists. Family practitioners recognized only 26% of all biopsied neoplastic and cystic skin lesions correctly, compared with 75%, recognized by dermatologists. Plastic surgeons were also considerably behind dermatologists, since they recognized correctly 45% of these lesions.

When examined separately, premalignant and malignant lesions were diagnosed correctly by dermatologists in 67% of cases, versus 11% and 44% diagnosed by family physicians and plastic surgeons, respectively. All the differences were statistically significant (5). A study from Australia showed that general practitioners agreed with dermatologist in 63% of cases in diagnosing benign nevi, but had a lower concordance rate in diagnosing melanomas and seborrheic keratoses (38% and 24%, respectively). Overall, dermatologists made correct diagnosis in 77% of cases prior to biopsy, and general practitioners made correct diagnosis in 24% of cases prior to biopsy. Diagnoses made by dermatologists, did not match the histology mostly in skin tumors and skin conditions where clinical diagnosis was not always possible or reliable. This extremely low level of agreement suggests a need for improving general practitioners' skills for recognizing dermatological conditions, especially in diagnosing skin cancer, through both undergraduate and postgraduate education (9). However, it is not determined whether further general practitioners' education will reduce the health-care costs. The increased costs, due to higher professional fees, is counterbalanced by greater laboratory costs, more misdiagnoses, subsequent return visits and/or referrals by family physicians.

Since the 1980s, there have been dedicated Pigmented Lesion Clinics (PLCs) in the UK, aimed to provide general practitioners (GPs) with a rapid referral system for lesions clinically suspected of being malignant melanomas (MMs). The time interval between GP referral and attendance at the clinic, diagnostic accuracy, and the time interval to definitive surgery, are the most important measures when comparing the efficacy of a referral system. The false negative rate (FNR), defined as the ratio between the number of false-negative clinical diagnoses to the number of histologically diagnosed MMs, represents a particularly important measure of diagnostic accuracy, being responsible for a significant delay in lesion excision. The accuracy of clinical diagnosis of MMs has been reported to increase with level of experience. Thus, it is not surprising that the FPR was significantly lower in Pigmented Lesion Clinics, staffed by dermatologists, particularly experienced in clinical diagnosis of pigmented lesions, than in General Dermatology and Plastic Surgery Clinics. Dermoscopy was not used in any of the clinics (10). However, when interobserver agreement of PLs malignancy risk was made by a pigmented lesion specialist (dermoscopy-experienced

dermatologist) and an artificial computer algorithm-based automated digital dermoscopy system, a low agreement was found. While digital dermoscopy confirmed its advantages, the computer-based algorithm requires further development and validation. From a clinical point of view, the main advantages of digital dermoscopy were high quality dermoscopic images, that allowed for enhanced patient education of clinical atypical nevi and melanoma warning signs (11).

In a study comparing self-reported and dermatologists' diagnoses, a low agreement was observed for five chronic skin diseases: acne, eczema, fungal infections, psoriasis and seborrheic dermatitis (12). Self-reports underestimated the actual prevalence of four of five diseases. Many cases of chronic skin diseases were diagnosed in patients who did not report them. Diagnostic knowledge was poorer in those above 54 years of age, and better in patients treated for the condition and those with impairment of social life. The results showed a need for dermatologists to explain to the population what these five diseases really are (12).

Case mix and diagnosis-related groups

The changing healthcare environment world-wide is leading to extensive use of so called "per case" payment systems. Currently, these systems are based on diagnosis-related groups. Diagnosis-related groups (DRGs) are primarily formed based on the patients primary diagnosis. Case mix refers to the number and types of patients treated, classified by diagnosis in the DRGs (13). It can be used for different purposes, but the main aim of DRG implementation was cost control, by setting hospital payments for all payers at a fixed DRG rate per admission. Thus, the future of inpatient dermatology is depiction of well-established diagnostic and treatment standards (14).

One of the main difficulties in dermatology is the problem how to offer highly specialized and extensive inpatient care. Patients with skin diseases remain difficult to categorize and to depict in patient classifications and case mix systems. The success of DRGs depends on the ability to distinguish severe and costly cases and less complex cases. If solutions are not found, dermatologists will be tempted to reduce their services or take action on early discharge and new admission (14). Many hospitals rely on other physicians, non-dermatologists, to care for dermatology inpatients. However, it should not be ignored that improvement in the quality of medical care will, and must be conceptually defined and implemented.

Conclusions

Since the ability of non-dermatologists, including internists and general practitioners, to make accurate diagnosis of some skin disorders, does not differ from that of undergraduate medical students, with no previous dermatologic experience, their diagnostic skills, especially in the area of skin cancers, need further improvement. Meanwhile, in order to minimize adverse treatment outcomes, and to decrease health-care costs, a low threshold for dermatology specialists referral should be encouraged. When comparing response rates for various modalities, one should bear in mind that the best interest of most inpatients with skin diseases is to be cared for by dermatologists in adequate clinical conditions.

References:

1. Cox NH. A literally blinded trial of palpation in dermatology diagnosis. *J Am Acad Dermatol* 2007;56:949-51.
2. Boulos KM. Map of dermatology: web image browser for differential diagnosis in dermatology. *Indian J Dermatol Venereol Leprol* 2006;72(1):72-4.
3. Janković A. Possibilities of teledermatology in marking clinical, dermoscopic and histologic diagnosis [dissertation]. Niš (Serbia): Medicinski fakultet; 2007.
4. Wallach D, Coste J, Tilles G, Taieb. The first images of atopic dermatitis: an attempt at retrospective diagnosis in dermatology. *J Am Acad Dermatol* 2005;53:684-9.
5. Sellheyer K, Bergfeld WE. A retrospective biopsy study of the clinical diagnostic accuracy of common skin diseases by different specialities compared with dermatology. *J Am Acad Dermatol* 2005;52:823-30.
6. Whited JD, Hall RP. Diagnostic accuracy and precision in assessing dermatologic diseases. *Arch Dermatol* 1997; 133:1409-15.
7. Chang CL, Chen CH. Applying decision tree and neural network to increase quality of dermatologic diagnosis. *Expert System Applications* 2009;36(2 Pt 2):4035-41.
8. Kulatunga-Moruzi C, Brooks LR. Coordination of analytic and similarity-based processing strategies and expertise in dermatological diagnosis. *Teach Learn Dermatol* 2001;13(2):110-6.
9. Tran H, Chen K, Lim AC, Jabbour J, Shumack S. Assessing diagnostic skill in dermatology: A comparison between general practitioners and dermatologists. *Aust J Dermatol* 2005;46:230-4.
10. Osborne FE, Chave TA, Hutchinson PE. Comparison of diagnostic accuracy for cutaneous malignant melanoma between general dermatology, plastic surgery and pigmented lesion clinics. *Br J Dermatol* 2003;148:252-8.
11. Boldrick JC, Layton CJ, Nguyen J, Swetter S. Evaluation of digital dermoscopy in a pigmented lesion clinic: clinician versus computer assessment of malignancy risk. *J Am Acad Dermatol* 2007;56:417-21.
12. Jagou M, Bastuji-Garin S, Bourdon-Lanoy E, Penso-Assathiany D, Roujeau JC. Poor agreement between self-reported and dermatologist's diagnosis for five common dermatoses. *Br J Dermatol* 2006;155:1006-12.
13. Roger France FH. Case mix use in 25 countries: a migration success but international comparisons failure. *Int J Med Inf* 2003;70:215-9.
14. Hensen P, Furstenberg T, Luger TA, Steinhoff M, Roeder N. Case mix measures and diagnosis-related groups: opportunities and threats for inpatient dermatology. *J Eur Acad Dermatol Venereol* 2005;19:582-8.