

Інформаційні технології в медицині

УДК 519:616-079.4:616.5

О.В. Висоцька¹, Е.М. Солошенко², А.І. Печерська¹

¹Харківський національний університет радіоелектроніки, Харків

²Інститут дерматології та венерології Академії медичних наук України, Харків

МЕТОД ВИЗНАЧЕННЯ ІНФОРМАТИВНИХ ОЗНАК ДЛЯ ДИФЕРЕНЦІАЛЬНОЇ ДІАГНОСТИКИ ПОШИРЕНИХ ДЕРМАТОЗІВ

В статті розроблено комплексний метод визначення інформативних критеріїв для диференціальної діагностики таких поширених дерматозів, як псоріаз, екзема, лікарська хвороба, atopічний дерматит та дерматит Дюринга. Розроблений метод поєднує формалізацію експертної інформації, визначення узгодженості думок експертів, розрахунок узагальненої експертної думки та інформаційного критерію.

Ключові слова: інформативні ознаки, диференціальна діагностика, поширені дерматози.

Вступ

Постановка проблеми. Багатофакторність формування механізмів розвитку захворювань шкіри створює трудомісткість діагностичної роботи. На практиці при визначенні диференціального діагнозу пацієнта з патологією шкіри лікарі зустрічаються з труднощами, пов'язаними з різноманіттям нозологічних форм та їхніх стадій, кожна з яких, по-перше, має свої особливості, а, по-друге, має велику кількість типових і іншим захворюванням чи стадіям ознак. При цьому у пацієнта може бути присутня й супутня патологія внутрішніх органів та систем, що ускладнює проведення диференціальної діагностики.

Такі поширені дерматози, як псоріаз, екзема, лікарська хвороба, atopічний дерматит, дерматит Дюринга дуже часто супроводжуються однаковими клінічними симптомами. В такому випадку для конкретизації патології хворого проводиться велика кількість лабораторних досліджень. Тому при верифікації діагнозу необхідно враховувати істотність симптому для того чи іншого дерматозу. Оцінити вплив кожного чинника, визначити його інформативність при діагностуванні поширених дерматозів та відібрати найбільш інформативні ознаки шляхом клінічного спостереження без допомоги математичних методів та інформаційних технологій дуже складно [1 – 2].

Аналіз останніх досліджень та публікацій. В процесі діагностування експерт-дерматолог користується механізмом асоціацій, заснованим на теоретичних знаннях та практичному досвіді. Асоціації виникають в процесі обстеження пацієнта – під час аналізу клінічних проявів захворювання та результатів досліджень. Тому, оцінюючи інформативність діагностичних ознак, необхідно в першу чергу вербалізувати та формалізувати знання експерта-дерматолога (яке не завжди в явній формі уявляє сам експерт), а вже потім використати його при ви-

значенні інформативних ознак [3]. При цьому припускається, що думка колегії експертів надійніша за думку окремого експерта.

Запорукою отримання якісної експертної інформації є вибір коректних способів її отримання, а також математичної обробки отриманих результатів та визначення узагальнених думок експертів.

Для визначення інформативних ознак використовуються наступні методи:

- статистичні методи;
- комбінаторні методи перебору різних варіантів поєднання ознак;
- лексикографічне впорядкування.

Доцільність використання того чи іншого методу визначається характером інформації, що обробляється [4]. Враховуючи те, що при диференціальній діагностиці поширених дерматозів використовуються не тільки якісні, а й кількісні ознаки, які мають різні одиниці виміру, доцільним є використання комплексу декількох методів.

Постановка завдання. Метою дослідження є розробка комплексного методу визначення інформативних ознак для диференціальної діагностики поширених дерматозів, що дозволить формалізувати експертні знання та виділити кортеж найбільш вагомих ознак і, як наслідок, підвищити якість диференціювання нозологічних одиниць зі схожими клінічними проявами.

Метод визначення інформативних ознак для диференціальної діагностики поширених дерматозів

Концепція методу визначення інформативних ознак для диференціальної діагностики поширених дерматозів полягає у визначенні інформативності всіх діагностичних ознак шкірної патології та відборі такого набору найбільш інформативних з них,

який дозволить проводити коректну диференціацію діагнозів, та передбачає:

- формування експертної групи;
- збір експертної інформації про дерматози;
- обробку експертної інформації;
- визначення інформативності діагностичних ознак.

На першому етапі роботи формується експертна група, тобто вирішується завдання кількісного та якісного відбору експертів.

Визначення чисельності групи експертів проводиться на основі «прагматичного» підходу, згідно якого для забезпечення якісної експертизи достатньо сформувати групу з 3 експертів.

Якісний відбір проводиться на основі доступної інформації про професійну підготовку кандидатів.

На другому етапі проводиться самооцінка експертів, тобто кожен експерт самостійно визначає рівень своєї компетентності у вигляді коефіцієнта самооцінки k_c ($0 < k_c < 10$). На основі k_c обчислюються нормовані коефіцієнти компетентності $k_{на}$ кожного експерта за формулою:

$$k_{на} = k_{са} / \sum_{a=1}^b k_{са}$$

де $k_{на}$ – нормований коефіцієнт компетентності a -го експерта; $k_{са}$ – коефіцієнт самооцінки a -го експерта; $a = 1, 2, \dots, b$ – порядковий номер експерта; b – кількість експертів [5].

Результати самооцінки та розрахунку коефіцієнтів компетентності обраних на першому етапі експертів наведено у табл. 1.

Таблиця 1
Коефіцієнти компетентності експертів за фахом «Дерматовенерологія»

Порядковий номер експерта, a	Коефіцієнти компетентності	
	Коефіцієнт самооцінки $k_{са}$	Нормований коефіцієнт компетентності $k_{на}$
1	10,00	0,38
2	8,00	0,31
3	8,00	0,31

Коефіцієнти компетентності використовуються для визначення узагальнених оцінок кожної з ознак шкірної патології на основі оцінок експертів.

На третьому етапі проводиться тестування, тобто збирається експертна інформація про захворювання, між якими проводиться диференційна діагностика, а саме, визначається комплекс, ступінь вираженості ознак, властивих кожному захворюванню.

Збір експертної інформації здійснюється на основі методу структуризації прийняття рішень. При цьому експертне опитування проводиться в декілька етапів, формуючи «дерево рішень». Кожний етап такого експертного опитування формує один етап «дерева рішень». В методі структуризації прийняття

рішень використовується дельфійська процедура обробки експертних відповідей.

На четвертому етапі проводиться числова ідентифікація лінгвістичних термів, а також проводиться ранжування ознак за частотою наявності при кожній патології. Наприклад, при оцінюванні імовірності присутності кожної з ознак при кожному з дерматозів, експерти визначили, що ознака може проявлятися «ніколи», «рідко», «інколи», «завжди» та «завжди». Кожному з цих лінгвістичних термів, що складають п'ятирівневу градацію, за допомогою вербально-графічного методу наводиться відповідна числова оцінка. Для того, щоб запобігти випадку переваги сумарного ефекту ознак з малою вагою над ознаками з великою вагою, рівні ієрархії відрізняються на порядок (рис. 1).

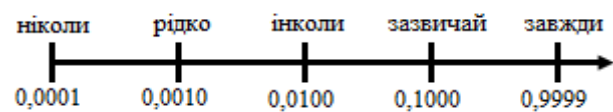


Рис. 1. Ідентифікація лінгвістичних термів, що виражають імовірність присутності ознаки при дерматопатології

Ці ідентифіковані лінгвістичні терми є імовірністю c_g , $g = 1, 2, \dots, G$. Таким чином, лінгвістичному терму «ніколи» відповідає імовірність $c_1 = 0,0001$; терму «рідко» – імовірність $c_2 = 0,0010$; терму «інколи» – імовірність $c_3 = 0,0100$; терму «завжди» – імовірність $c_4 = 0,1000$; терму «завжди» – імовірність $c_5 = 0,9999$.

При оцінюванні ступеня прояву ознаки, експерти визначили, що прояв якісної ознаки може бути «відсутній», «слабкий», «помірний», «сильний» та «дуже сильний». Ідентифікація цих лінгвістичних термів також проводиться за допомогою вербально-графічного методу, проте не потребує значного відділення числових рівнів ієрархії один від одного (рис. 2).



Рис. 2. Ідентифікація лінгвістичних термів, що виражають ступінь прояву ознаки

Для кількісного позначення ступеня прояву ознаки користуються величиною вираженості O_s , $s = 1, 2, \dots, S$. Таким чином, лінгвістичному терму «прояв ознаки відсутній» відповідає вираженість $O_1 = 0,1$; терму «прояв ознаки слабкий» – вираженість $O_2 = 0,3$; терму «прояв ознаки помірний» – вираженість $O_3 = 0,5$; терму «прояв ознаки сильний» – вираженість $O_4 = 0,7$; терму «прояв ознаки дуже сильний» – вираженість $O_5 = 0,9$.

Експертна інформація, приведена до єдиної числової форми, $h_{ас}$, заноситься до таблиць. Наприклад, експертні оцінки імовірності прояву ознаки

«підвищення вмісту у сироватці крові Т-лімфоцитів (СДЗ)» при псоріазі наведені у табл. 2.

Таблиця 2

Експертні оцінки h_{as} імовірності прояву ознаки «підвищення вмісту у сироватці крові Т-лімфоцитів (СДЗ)» при псоріазі

Порядковий номер експерта	Експертні оцінки h_{as} кожної вираженості ознаки, O_s				
	0,1	0,3	0,5	0,7	0,9
1	0,0001	0,0010	0,0100	0,9999	0,1000
2	0,0001	0,0010	0,0100	0,9999	0,1000
3	0,0001	0,0010	0,1000	0,9999	0,0100

На четвертому етапі оцінюється ступінь узгодженості знань експертів. Для цього розраховується ентропійний коефіцієнт конкордації (коефіцієнт згоди) за формулою:

$$W_E = 1 - \frac{H}{H_{\max}}, \quad (1)$$

де H – ентропія, яку розраховують за формулою:

$$H = - \sum_{s=1}^S \sum_{g=1}^G p_{sg} \log p_{sg}, \quad (2)$$

а H_{\max} – максимальне значення ентропії.

У формулі для ентропії p_{sg} – оцінки вірогідностей c_g -ї імовірності, наданої s -й вираженості ознаки O_s . Ці оцінки вірогідностей розраховуються у вигляді відношення кількості експертів m_{sg} , які надали вираженості ознаки O_s імовірність c_g , до загальної кількості експертів:

$$p_{sg} = \frac{m_{sg}}{b}. \quad (3)$$

Максимальне значення ентропії досягається при рівновірогідному розподілі імовірностей, тобто коли $m_{sg} = b/s$. Тоді

$$p_{sg} = \frac{b}{bS} = \frac{1}{S}.$$

Виходячи з цього, максимальна ентропія визначається за формулою:

$$H_{\max} = S \log S. \quad (4)$$

Коефіцієнт згоди змінюється від нуля до одиниці. При $W_E = 0$ згода експертів відсутня, а при $W_E = 1$ думки експертів повністю збігаються. Вважається, що при значенні ентропійного коефіцієнта конкордації $W_E \geq 0,75$, узгодження думок експертів є достатнім [6 – 7].

Наприклад, визначимо узгодженість думок експертів щодо імовірності ознаки «підвищення вмісту у сироватці крові Т-лімфоцитів (СДЗ)» при псоріазі.

Підрахуємо у табл. 2 кількість експертів m_{sg} , які надали вираженості ознаки O_s імовірність c_g . Результати підрахунку заносимо у табл. 3.

Розраховані за формулою (3) значення оцінок вірогідностей заносимо до табл. 4.

Таблиця 3

Результати визначення m_{sg}

Імовірність прояву ознаки, c_g .	Кількість експертів m_{sg} , які надали вираженості ознаки O_s імовірність c_g				
	0,1	0,3	0,5	0,7	0,9
0,0001	0	0	0	3	0
0,0010	0	0	1	0	2
0,0100	0	0	2	0	1
0,1000	0	3	0	0	0
0,9999	3	0	0	0	0

Таблиця 4

Значення оцінок вірогідностей імовірності ознаки «підвищення вмісту у сироватці крові Т-лімфоцитів (СДЗ)» при псоріазі

Імовірність, c_g	Значення оцінок вірогідностей імовірності ознаки p_{sg} для кожної вираженості ознаки O_s				
	0,1	0,3	0,5	0,7	0,9
0,0001	0	0	0	1	0
0,0010	0	0	0,33	0	0,67
0,0100	0	0	0,67	0	0,33
0,1000	0	1	0	0	0
0,9999	1	0	0	0	0

За формулою (2) розраховуємо значення ентропії. Для даних, наведених у табл. 4, $H=0,55$.

За формулою (4) розраховуємо значення максимальної ентропії. У нашому прикладі $H_{\max}=3,49$

За формулою (1) розраховуємо значення ентропійного коефіцієнта конкордації. У нашому прикладі $W_E=0,84$, тобто розраховане значення коефіцієнта згоди вище порогового значення, тобто думки експертів узгоджені. В іншому випадку експерти мають переглянути свої думки.

На п'ятому етапі, якщо оцінки експертів не узгоджені, тобто коефіцієнт згоди менше 0,75, результати обробки експертної інформації повертаються експертам та проводиться повторне опитування.

На шостому етапі, якщо оцінки експертів узгоджені, тобто коефіцієнт згоди більше 0,75, проводиться визначення узагальнених експертних оцінок. Для цього доцільно скористатися наступною формулою:

$$h_s = \sum_{a=1}^b h_{as} \cdot k_{на}, \quad (5)$$

де $k_{на}$ – нормований коефіцієнт компетентності a -го експерта, наведений у табл. 1; $a = 1, 2, \dots, b$ – порядковий номер експерта; b – кількість експертів; h_{as} – експертні оцінки імовірності прояву ознаки «підвищення вмісту у сироватці крові Т-лімфоцитів (СДЗ)» при псоріазі, наведені у табл. 2.

Результати розрахунку узагальненої імовірності прояву ознаки «підвищення вмісту у сироватці крові Т-лімфоцитів (СДЗ)» при псоріазі наведені у табл. 5.

На п'ятому етапі проводиться оцінка інформативності діагностичних ознак.

Таблиця 5

Узагальнені експертні оцінки h_s імовірності прояву ознаки O_s «підвищення вмісту у сироватці крові Т-лімфоцитів (СД3)» при псоріазі

O_s	0,1	0,3	0,5	0,7	0,9
h_s	0,0001	0,001	0,0379	1,0000	0,0721

Оскільки при діагностуванні експерти користуються не тільки якісними, а й кількісними ознаками, які мають різну природу (різні одиниці, інтервали і шкали виміру), тобто, показники незрівняні між собою, то для їх ідентифікації необхідно використовувати інформаційний критерій, що не залежить від одиниць виміру ознак. Даним вимогам відповідає критерій інформативності, що володіє властивостями ентропії, який враховує лише відносні значення розподілу ознак на елементах класів розпізнавання [14].

З математичної точки зору всі пацієнти розглядаються як сукупність об'єктів із змінними кількісними і якісними характеристиками.

Нехай X матриця з позитивними елементами

$$X \in R_+^{n \times m} = \{X = (x_{ij}) \mid x_{ij} > 0, i = \overline{1, n}; j = \overline{1, m}\},$$

що має n рядків і m стовпців. Елементами матриці є значення ознак дерматологічного статусу кожного пацієнта.

Також нехай k – число розпізнаваних класів (діагнозів); $\xi(x_{ij})$ – номер класу, якому відповідає значення x_{ij} j -ї ознаки у i -го пацієнта і нехай

$$p \in \Delta_n = \{p = (p_1, p_2, \dots, p_n)^T \mid \sum_{i=1}^n p_i = 1, p_i \geq 0, \forall i = \overline{1, n}\}.$$

Тоді інформативність j -ї ознаки (стовпця X_j матриці X) можна оцінити на основі рішення завдання формування інформаційного критерію:

$$I(x) = h(x) - h(x/\Omega) \quad (6)$$

$$I(x) = \ln(p^T x \cdot x^{-T} p) - \min_{w \in \Omega} \ln((p/w)^{-T} x \cdot x^{-T} (p \cdot w))$$

де $h(x) = \ln(x^{-T} p \cdot p^T x) = \ln \left(\sum_{i=1}^n \frac{p_i}{x_i} \cdot \sum_{i=1}^n p_i x_i \right)$ з області

визначення Ω у вигляді:

$$\Omega = \{w \in R_+^{n \times 1} \mid w_u = w_v \Leftrightarrow \xi(x_{ui}) = \xi(x_{vi}) \forall j = \overline{1, m}\}.$$

Інформативність ознаки тим вища, чим ближче значення інформаційного критерію до одиниці [8 – 9]. Вважається, що при значенні $I(x_j) \geq 0,8$ критерій x_j є інформативним.

Для визначення діагностичних показників, інформативних при диференціації дерматозів, була досліджена група з 450 хворих на поширені дерматози, а саме 150 хворих на псоріаз, 100 – на екзему, 80 – на лікарську хворобу, 70 – на atopічний дерматит та 50 – на дерматит Дюринга. Кожний пацієнт був оглянутий трьома експертами та пройшов лабораторні дослідження, в результаті чого встановлено

коректний диференційний діагноз та проведено ефективне лікування. Стан кожного хворого характеризувався 94 кількісними лабораторними показниками (показники гемограми, імунологічного, алергологічного, біохімічного та загальноклінічного досліджень) та 23 якісними показниками, отриманими в результаті опитування та огляду хворого. Всі ознаки були закодовані й наведені у відповідність 117-мірному вектору, який враховує наявність та вираженість кожної ознаки. Відповідно до діагнозу, пацієнти були розподілені на 5 груп.

В результаті розрахунку інформативності кожної з ознак за формулою (6) з початкових 117 ознак вдалося визначити 14 інформативних (рис. 3), які адекватно віддзеркалюють патологічні процеси при різних дерматозах.

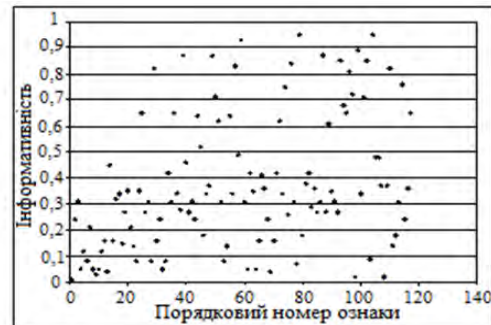


Рис. 3. Діаграма розподілу інформативності діагностичних ознак

Результати класифікації хворих за методом, описаним в [10], з використанням виділених інформативних показників наведено в табл. 6. Аналіз результатів класифікації свідчить, що прогноз для груп 1 (псоріаз), 2 (екзема), 3 (лікарська хвороба), 4 (атопічний дерматит), 5 (дерматит Дюринга) зроблено вірно, та коректно класифіковано, в загальній складності, 96,4% усіх хворих на дерматози. Відповідно, точність диференціальної діагностики поширених дерматозів для обстежених хворих може досягати 96,4% в тому випадку, якщо вибірка буде ідентична тим пацієнтам, дані котрих послужили основою для прогнозу.

Таким чином, диференціальну діагностику проаналізованих п'яти дерматозів можна проводити за допомогою наступних показників: вміст у сироватці крові Т-лімфоцитів (СД3), вміст у сироватці крові Т-лімфоцитів хелперів (СД4), вміст у сироватці крові Т-цитотоксичних клітин (СД 8); вміст у сироватці крові фактору апоптозу (FASx); вміст у сироватці крові поліпептиду ендотелеїну-1; вміст у сироватці крові вітаміну Е, вміст у сироватці крові інтерлейкіну 6 (ИЛ 6), вміст у сироватці крові інтерлейкіну 8 (ИЛ8), вміст у сироватці крові інтерлейкіну 10 (ИЛ 10), вміст у сироватці крові фактору некрозу пухлини, вміст у сироватці крові молекул середньої маси (МСМ), вміст у сироватці крові імуноглобуліну А (Ig А); вміст у сироватці крові імуноглобуліну G (Ig G), вміст у сироватці крові загального імуноглобуліну Е (Ig E).

Результати класифікації

	Діагноз	Прогнозована приналежність до групи					Всього	
		Група 1	Група 2	Група 3	Група 4	Група 5		
Реальна приналежність до групи	Кількість	Група 1	147	2	0	1	0	150
		Група 2	6	92	1	1	0	100
		Група 3	0	3	76	0	1	80
		Група 4	0	1	2	66	1	70
		Група 5	1	0	0	1	48	50
	%	Група 1	98	1	0	1	0	100
		Група 2	6	92	1	1	0	100
		Група 3	0	4	95	0	1	100
		Група 4	0	1	3	94	1	100
		Група 5	0	0	0	4	96	100

Висновки

Завдяки поєднанню різних методів і підходів оброблена та проаналізована велика кількість експертної діагностичної інформації різного характеру. Згідно прагматичного підходу сформована експертна група з трьох експертів. Проведено експертне опитування на основі методу структуризації прийняття рішень. За допомогою вербально-графічного методу ідентифіковані лінгвістичні терми експертів. На основі визначення ентропійного коефіцієнту конкордації перевірена узгодженість думок експертів. В результаті розрахунку інформаційного критерію визначено 14 інформативних ознак, які дозволяють диференціювати такі поширені дерматози, як псоріаз, екзема, лікарська хвороба, атопічний дерматит та дерматит Дюринга.

Запропонований метод дозволяє удосконалити диференціацію захворювань шкіри з однаковими клінічними ознаками, що є передумовою постановки коректного диференційного діагнозу та вибору адекватної лікувальної тактики.

Список літератури

1. Мавров І.І. Дерматологія і венерологія в контексті загальної медицини // *Доктор*. – 2005. – № 1. – С. 4-8.
2. Туркевич Ю.М. Шкірні та венеричні хвороби / Ю.М. Туркевич, О.Ю. Туркевич, О.О. Сизон. – Львів: ЗУКЦ, 2011. – 239 с.

3. Кобринский Б.А. Возможность вскрытия интуитивных представлений врачей при групповом извлечении знаний / Б.А. Кобринский // Десятая национальная конференция по искусственному интеллекту с международным участием КИИ-2006 (25-28 сентября 2006 г., Обнинск): Труды конференции. В 3-т. Т.2. – М.: Физматлит, 2006. – С. 653-659.

4. Котов Ю.Б. Новые математические подходы к задачам медицинской диагностики; Изд.2 / Ю.Б. Котов. – М.: URSS, 2011. – 328 с.

5. Трахтенгерц Э.А. Компьютерная поддержка принятия решений / Э.А. Трахтенгерц. – М.: СИНТЕГ, 1998. – 237 с.

6. Алексеев С.И. Исследование систем управления / С.И. Алексеев. – М.: Изд. центр ЕАОИ, 2008. – 195 с.

7. Малин А.С. Исследование систем управления: учебн. для вузов / А.С. Малин, В.И. Мухин. – М.: ГУ 2002. – 400 с.

8. Куренков Н.И. Информационный критерий и его использование для решения задач обработки многомерных данных / Н.И. Куренков, С.Н. Ананьев // Информационные технологии. – 2007. – № 9. – С. 59-64.

9. Печерская А.И. Определение информативности диагностических параметров для дифференциации распространенных дерматозов / А.И. Печерская // Прикладная радиоэлектроника. – X.: ХНУРЕ, 2008. – Том 7, №4. – С. 356-361.

10. Печерская А.И. Применение метода многопараметрического выбора оптимального решения в дифференциальной диагностике дерматопатологии / А.И. Печерская // Матеріали XI Міжнародної конференції з біоніки та прикладної біофізики. – К.: НТУ «КПІ», 2010. – С. 68-69.

Надійшла до редколегії 18.05.2012

Рецензент: д-р фіз.-мат. наук, проф. А.І. Бих, Харківський національний університет радіоелектроніки, Харків.

МЕТОД ОПРЕДЕЛЕНИЯ ИНФОРМАТИВНЫХ ПРИЗНАКОВ ДЛЯ ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ РАСПРОСТРАНЕННЫХ ДЕРМАТОЗОВ

Е.В. Высоцкая, Э.Н. Солошенко, А.И. Печерская

В статье разработан комплексный метод определения информативных признаков для дифференциальной диагностики таких распространенных дерматозов как псориаз, экзема, лекарственная болезнь, атопический дерматит и дерматит Дюринга. Разработанный метод объединяет формализацию экспертной информации, определение согласованности мнений экспертов, расчет обобщенного экспертного мнения и информационного критерия.

Ключевые слова: информативные признаки, дифференциальная диагностика, распространенные дерматозы.

THE METHOD FOR DEFINITION OF INFORMATIVE PARAMETERS FOR DIFFERENTIAL DIAGNOSTICS WIDELY-SPREAD DERMATOSIS

O.V. Vysotska, E.M. Soloshenko, A.I. Pecherska

The complex method, which allows to define informative parameters for differential diagnostics of such widely-spread dermatosis as psoriasis, eczema, drug disease, atopic dermatitis, dermatitis herpetiformis Dühring, is created in the article. Formalization of expert information, definition of consistency the expert's considerations, calculation of the generalized expert's consideration and information criterion are incorporated by created method.

Keywords: informing signs, differential diagnostics, widespread dermatosis.