

ОЦЕНКА КОЛИЧЕСТВА ДИАГНОСТИЧЕСКИХ ПАРАМЕТРОВ ДЛЯ ДИФФЕРЕНЦИАЦИИ ЗАБОЛЕВАНИЙ ЛЁГКИХ

А.М. Шайдук^{1,2}, И.В. Молодкин²

¹ Алтайский государственный медицинский университет, Барнаул

² Алтайский государственный университет, Барнаул

Рассмотрена методика оценки необходимого числа диагностических параметров для диагностики некоторых заболеваний лёгких. Диагностические параметры, полученные анализом медицинских изображений рентгеновской компьютерной томографии (РКТ) внутренних органов, имеют статистическую зависимость от типа патологии. Это даёт возможность статистической дифференциации патологий. Нет единственного диагностического признака, который бы однозначно определял тип заболевания, только совокупность таких параметров даёт возможность вероятностной диагностики. Поэтому определение количества необходимых параметров для постановки достоверного диагноза является актуальной задачей. В работе представлена методика оценки количества таких параметров на примере двух заболеваний. Проведен количественный расчет для дифференциации патологий “рак – туберкулез” на основе анализа диагностических параметров больных с верифицированными диагнозами.

Ключевые слова: шаровидные образования в лёгких, компьютерная томография, вероятностная диагностика, плотность Хаунсфилда

Введение

При постановке диагноза врач оперирует некими качественными и количественными характеристиками пациента, делая выводы на основе своих знаний и опыта [1, 2]. Например, врач-диагност, рассматривая область возможной патологии на рентгеновском снимке, может учитывать оптическую плотность изображения, стандартное отклонение этой плотности, фрактальную размерность области, степень сложности контура патологии и другие диагностические признаки [3–5]. Далее будем рассматривать только те диагностические признаки, которые можно выразить количественно и будем называть их диагностическими параметрами.

В медицинской практике очень часто для разделения типа заболеваний требуется несколько параметров, так как параметра, одно-

значно определяющего конкретное заболевание, не существует. Решение приходится принимать на основе совокупности диагностических параметров. При этом каждое конкретное значение диагностического параметра может реализовываться при различных заболеваниях. Но все же вероятность реализации конкретного значения диагностического параметра для разных заболеваний различна [6], что и позволяет экспертам с некоторой вероятностью определять тип заболевания по совокупности диагностических параметров (рис. 1).

В этой связи возникает вопрос о количестве требуемых параметров в этой совокупности. Сколько диагностических параметров нужно знать, чтобы установить тип заболевания с требуемой вероятностью? Как зависит число диагностических параметров от их статистических свойств?

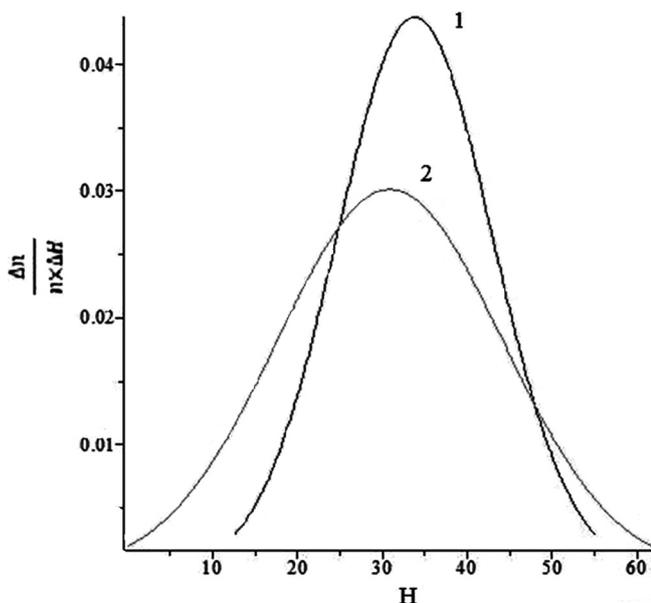


Рис. 1. Аппроксимированные распределения плотности Хаунсфилда для больных раком лёгкого (1) и туберкулёза (2)

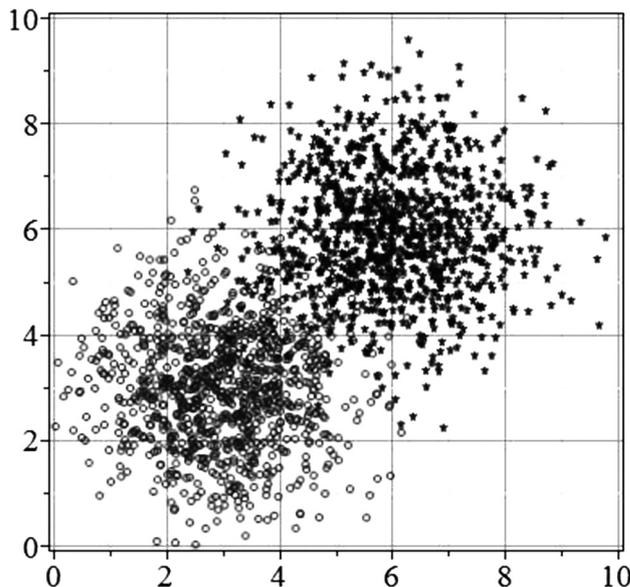


Рис. 2. Одновременное изображение пациентов с заболеваниями А и В в многомерном пространстве симптомокомплекса

Методология

Рассмотрим задачу разделения двух заболеваний по нескольким диагностическим параметрам. Пусть мы имеем набор диагностических параметров x_1, x_2, x_3, \dots , характеризующих область интереса или пациента. Пусть для заболевания А и каждого из диагностических параметров известны его функции распределения (гистограммы) $f_{A1}(x_1), f_{A2}(x_2), f_{A3}(x_3) \dots$ для заболевания В также известны все функции распределения $f_{B1}(x_1), f_{B2}(x_2), f_{B3}(x_3) \dots$. Пусть для некоторого пациента мы можем получить конкретные значения этих диагностических параметров $x_1=x_{10}, x_2=x_{20}, x_3=x_{30} \dots$. Сколько диагностических параметров необходимо иметь, чтобы с вероятностью $P_{\text{крит}}$ выбрать заболевание А или В?

Введём абстрактное n -мерное пространство диагностических параметров. Совокупность диагностических параметров одного пациента будет изображаться точкой в этом пространстве [7]. Генеральная совокупность пациентов с одинаковым заболеванием А будет порождать в этом пространстве n -мерное облако своих диагностических параметров. Облака заболеваний А и В, к сожалению, будут иметь области пересечения (рис. 2).

В нашем случае, для постановки диагноза необходимо определить к какому из облаков (А или В) принадлежит n -мерная точка, характеризующая нового пациента. Из теории вероятности можно получить, что плотность облака А есть:

$$\rho_A(\bar{x}_A) = N f_{A1}(x_1) f_{A2}(x_2) \times \dots \times f_{An}(x_n). \tag{1}$$

Аналогично для В:

$$\rho_B(\bar{x}_B) = N f_{B1}(x_1) f_{B2}(x_2) \times \dots \times f_{Bn}(x_n),$$

где N – число пациентов в генеральной совокупности заболевания. Для нахождения вероятности заболевания А необходимо воспользоваться классическим её определением.

$$\begin{aligned} dN_A &= \rho_A(\bar{x}) d\Gamma \\ dN_B &= \rho_B(\bar{x}) d\Gamma \Rightarrow P_A = \frac{dN_A}{dN} = \frac{\rho_A(\bar{x})}{\rho_A(\bar{x}) + \rho_B(\bar{x})} \end{aligned} \tag{2}$$

$$dN = dN_A + dN_B$$

где $d\Gamma = dx_1 dx_2 dx_3 \dots dx_n$. Аналогично для В:

$$P_B = \frac{\rho_B(\bar{x})}{\rho_A(\bar{x}) + \rho_B(\bar{x})}.$$

Для пациента с параметрами $x_{10}, x_{20}, x_{30} \dots$ вероятность заболевания А будет равна:

$$P_A = \frac{\rho_A(x_{10}, x_{20}, x_{30}, \dots)}{\rho_A(x_{10}, x_{20}, x_{30}, \dots) + \rho_B(x_{10}, x_{20}, x_{30}, \dots)}. \tag{3}$$

Для нового пациента вероятность того, что при данном наборе диагностических параметров пациент страдает заболеванием А, дается соотношением (3) (формула Байеса). Таким образом, из соотношений (1–3) получим условия постановки диагноза заболевания А:

$$P_A = \frac{\prod_{i=1}^n f_{Ai}(x_{i0})}{\prod_{i=1}^n f_{Ai}(x_{i0}) + \prod_{i=1}^n f_{Bi}(x_{i0})} \geq P_{\text{крит}}. \quad (4)$$

В данном выражении число сомножителей n требуется увеличивать до тех пор, пока не выполнится условия для $P_{\text{крит}}$. Если оно не выполняется при любых n , указанное заболевание отсутствует. Это условие и определяет число n необходимых диагностических параметров.

Для дальнейшей оценки n введём коэффициент эффективности диагностического параметра, определяемого формулой (5).

$$K_{AB} = f_{Ai}(x_{i0}) / f_{Bi}(x_{i0}). \quad (5)$$

Из (4) и (5) находим

$$P_A = \frac{\prod_{i=1}^n K_{AB}(x_{i0})}{\prod_{i=1}^n K_{AB}(x_{i0}) + 1} \geq P_{\text{крит}}. \quad (6)$$

Откуда

$$\prod_{i=1}^n K_{AB}(x_{i0}) \geq \frac{P_{\text{крит}}}{1 - P_{\text{крит}}}. \quad (7)$$

Оценивая количество параметров, необходимых для постановки диагноза, предположим, что коэффициенты эффективности для всех диагностических параметров будет иметь одинаковое значение и окажутся равными, тогда выражение (7) преобразуется в степенную функцию:

$$\prod_{i=1}^n K_{AB}(x_{i0}) \approx K^n. \quad (8)$$

По нашим оценкам [8], для выбора типа шаровидного образования между патологиями “рак – туберкулез” этот коэффициент может находиться в диапазоне $1,1 < K < 1,5$. С учётом (8) мы можем найти количество необходимых параметров n :

$$n \geq (\ln(P_{\text{крит}}) - \ln(1 - P_{\text{крит}})) / \ln(K). \quad (9)$$

На рис. 3 представлен график зависимости числа диагностических параметров от характерного коэффициента эффективности K . Разумеется, значения n являются целыми числами и на графике они соединены непрерыв-

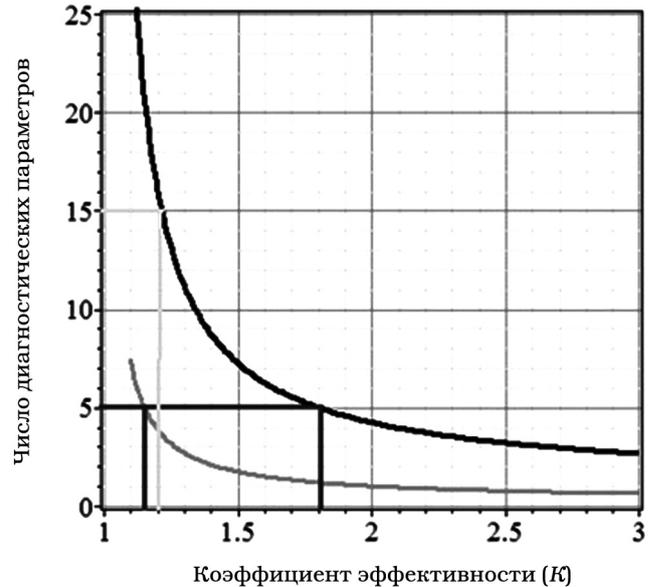


Рис. 3. Зависимость числа диагностических параметров от эффективности диагностических признаков.

ной кривой для наглядности. Верхний график соответствует $P_{\text{крит}} = 0,95$, нижний график построен для низкой степени уверенности в диагнозе $P_{\text{крит}} = 0,6$. Как видно из рис. 3, при небольших значениях коэффициента эффективности число требуемых диагностических параметров становится очень большим. Так, для $K=1,2$ это число уже достигает 15 диагностических параметров. Вряд ли врач-диагност способен различать совокупности, состоящие из 15 различных элементов, без помощи информационно-математических систем поддержки решений врача.

Приемлемым для “ручной” диагностики число диагностических параметров (на графике оно условно принято равным 5) становится лишь при наличии очень эффективных диагностических параметров (с коэффициентом эффективности $K=1,8$ и более). Если такие эффективные диагностические параметры отсутствуют, то врач ставит диагноз с невысокой надежностью (нижняя кривая на рис. 3)).

Результаты

Таким образом, по крайней мере для шаровидных образований в легких можно утверждать следующее.

- ✓ К сожалению, для успешной постановки диагноза при отсутствии высокоэффективных

признаков требуется слишком большое число диагностических параметров. Для постановки диагноза требуются информационно-математические системы поддержки решений врача, причем не как способ облегчить работу врача, а как математическая неизбежность.

- ✓ Без компьютерных систем поддержки решений врача нельзя с высокой степенью надежности поставить диагноз для “трудного” пациента.
- ✓ При малом числе малоэффективных диагностических параметров (которое только и может удержать в поле своего внимания врач) диагноз неизбежно ставится с невысокой надежностью.
- ✓ Для “ручной” постановки диагноза необходимы высокоэффективные диагностические параметры с коэффициентом эффективности $K > 1,8$, поиск которых всегда является в диагностике актуальным.

Список литературы

1. Коновалов В.К., Лобанов М.Н., Леонов С.Л., Шайдук А.М и соав. Денситометрия шаровидных образований легких с использованием искусственного нейрона. // Вестник Алтайской науки, 2013, № 2–1, С. 140 – 149.
2. Коновалов В.К., Лобанов М.Н., Леонов С.Л., Шайдук А.М и соав. Способ количественной оценки структуры шаровидных образований легких при мультиспиральной компьютерной томографии. // Вестник Алтайской науки, 2013, № 2–1, С. 149–152.
3. Коновалов В.К., Колмогоров В.Г., Лобанов М.Н. и соавт. Особенности селективной денситометрии шаровидных образований легких при мультиспиральной компьютерной томографии. // Росс. онкол. журнал, 2014, № 19(5), С. 14–23.
4. Коновалов В.К., Колмогоров В.Г., Лобанов М.Н. и соавт. Объемная денситометрическая диагностика шаровидных образований легких с использованием искусственного нейрона. // Росс. онкол. журнал, 2014, № 19(1), С. 25–31.
5. Коновалов, В.К., Шойхет, Я.Н., Федоров, В.В. и соавт. Прицельная 3d-реконструкция при изучении качественных характеристик поверхности шаровидных образований легких. // Проблемы клинической медицины, 2011, № 3–4, С. 20–25.
6. Молодкин И.В., Шайдук А.М., Коновалов В.К., Леонов С.Л. Статистические свойства диагностических параметров при денситометрии шаровидных образований в легком. // Известия Алтайского государственного университета, 2014, № 1–1(81), С. 218–222.
7. Ландау Л. Д., Лифшиц Е. Статистическая физика. – М.: Ч. I. 5-е изд., 1995 (2001, 2002).
8. Молодкин И.В., Леонов С.Л., Шайдук А.М., Коновалов В.К. Статистический анализ влияния типа патологии на количественные характеристики медицинских изображений. // Мед. физика, 2014, № 3(63), С. 43–47.

ESTIMATE OF THE NUMBER OF DIAGNOSTIC PARAMETERS FOR STATISTICAL DIFFERENTIATION LUNG DISEASE

A.M. Shayduk^{1,2}, I.V. Molodkin²

¹ Altai State Medical University, Barnaul, Russia

² Altai State University, Barnaul, Russia

The article discusses the method of estimating the required number of diagnostic parameters for the diagnosis of certain diseases of the lungs. Diagnostic parameters obtained by analyzing medical imaging X-ray computed tomography (CT) of the internal organs, are statistically dependent on the type of pathology. This enables statistical differentiation abnormalities. There is no single diagnostic feature that would uniquely determine the type of the disease, only a combination of these parameters enables the probabilistic diagnosis. Therefore, the determination of the number of parameters required for setting a definite diagnosis is an urgent task. The paper presents a method of estimating the number of parameters in the example of the two diseases. A quantitative calculation to differentiate pathologies “cancer – TB” on the basis of the analysis of diagnostic parameters of patients with verified diagnosis.

Key words: *spherical formation in the lungs, CT, diagnostics probability, density Hounsfield*

E-mail: tian.08@mail.ru