Dottorato di Ricerca in Biopatologia
XXIV Ciclo
Università degli Studi di Palermo

Analysis of a database to predict the result of allergy testing *in vivo* in patients with chronic nasal symptoms and the development of the software ARSTAT©

Tesi di dottorato della
Dott.ssa Maria Stefania Leto Barone

Coordinatore del Dottorato
Ch.mo Prof. Calogero Caruso

Tutor della Dottoranda
Ch.mo Prof. Gabriele Di Lorenzo
MED/09

UNIVERSITÀ DEGLI STUDI DI PALERMO

Dottorato di Ricerca in Biopatologia

ESTRATTO VERBALE DEL COLLEGIO DEI DOCENTI DEL DOTTORATO DI RICERCA IN BIOPATOLOGIA DEL 25 NOVEMBRE 2013

Il giorno 25 novembre 2013 alle ore 12.30 presso l'Aula del Porticato - Sezione di Patologia generale del Dipartimento di Biopatologia e Biotecnologie Mediche e Forensi si riunisce il collegio dei Docenti, giusta convocazione del Coordinatore, Prof. Calogero Caruso. Sono presenti i Docenti Professori: Calogero Caruso, Giuseppina Colonna Romano, Gabriele Di Lorenzo, Giuseppina Candore, Mario Mirisola, Vito Franco, Francesco Gervasi. Assente giustificata la prof. Stefania Grimaudo .Il Coordinatore constatato che la seduta è regolarmente costituita, invita la Prof.ssa Giuseppina Colonna-Romano a svolgere le funzioni di segretario e invita i dottorandi ad esporre il risultato delle loro ricerche.

OMISSIS

Valutazione annuale e finale dei dottorandi del XXIV ciclo (Relazione finale)

La Dott.ssa Maria Stefania Leto-Barone (Tutor prof. G. Di Lorenzo) presenta la relazione dell'attività svolta per la preparazione della tesi di dottorato "Analysis of a database to predict the result of allergy testing *in vivo* in patients with chronic nasal symptoms." ed espone quanto di seguito: i dati che si riferiscono a 1359 pazienti con diagnosi di rinite, visitati presso l'Ambulatorio di Allergologia del Dipartimento BioMedico di Medicina Interna e Specialistica (Di.Bi.M.I.S.) che fa parte del Dipartimento Assistenziale di Medicina Interna e Malattie Infettive dell'AOU Policlinico "Paolo Giaccone" di Palermo. I dati sono stati ricavati dalle cartelle dell'ambulatorio e sono stati tabulati e quindi analizzati con test statistici in base al tipo di variabile esaminata. Tutte le variabili risultate statisticamente significative, dal confronto di pazienti con test allergologico cutaneo positivo e negativo [skin prick test (SPT+) e (SPT-)], sono state esaminate con modelli di regressione logistica. Il migliore modello è stato verificato usando l'area sotto la curva con receiver operating characteristic (ROC). Quindi si è così elaborato un algoritmo tarato su un cutoff > 70 (SPT+) e < 70 (SPT-). Questo algoritmo è stato validato presso l'Ambulatorio di Allergologia dell'Ospedale Sandro Pertini di Roma. I pazienti erano inviati dai Medici di Medicina Generale , per eseguire lo SPT, perché affetti da una forma di rinite cronica. I pazienti erano inviati dai Medici di Medicina Generale, per eseguire lo SPT, perché affetti da una forma di rinite cronica. Questo Centro utilizzando il nostro questionario, ha identificato correttamente il 96% dei pazienti con SPT+ e il 94% dei pazienti con SPT-. Questi risultati hanno portato a intraprendere una collaborazione il Dipartimento di Scienze Economiche,Aziendali e Statistiche, dell'Università degli Studi di Palermo, con cui è stato elaborato un software che potrà essere utilizzato dal Medico di Medicina Generale, per valutare il miglior percorso diagnostico da intraprendere per i pazienti con sintomi nasali cronici.

UNIVERSITÀ DEGLI STUDI DI PALERMO · DIPARTIMENTO DI BIOPATOLOGIA E BIOTECNOLOGIE MEDICHE E FORENSI

Nel corso di questi anni l'intensa attività di ricerca ha permesso alla dott.ssa Leto Barone di produrre lavori scientifici, pubblicati su riviste sottoposte a valutazione ISI ed a severo controllo redazionale. Il collegio dei docenti è pertanto lieto di attestare il buon livello di preparazione acquisita e la valuta meritevole di conseguire il titolo di Dottore di Ricerca in Biopatologia.

OMISSIS

redatto, letto ed approvato seduta stante. La seduta è tolta alle ore 14.00

Il Segretario — Il Coordinatore

Prof. Giuseppina Colonna-Romano — Prof. Calogero Caruso

COPIA CONFORME ALL'ORIGINALE

UNIVERSITÀ DEGLI STUDI DI PALERMO · DIPARTIMENTO DI BIOPATOLOGIA E BIOTECNOLOGIE MEDICHE E FORENSI

Indice

Contents

Summary

Background

This thesis consists of 4 parts: (i) Introduction in which we present the clinical problem of rhinitis; (ii) the methods to evaluate the diagnostic choices; (iii) the rational errors in Allergy; (iv) the experimental part of thesis with which we developed the software ARSTAT, which is the application of the analysis reported.

Objective

We studied the ability of the logistic regression model, obtained by the evaluation of a database, to detect patients with positive allergy skin prick test (SPT) and patients with negative SPT. The model developed was validated using the data set obtained from another medical institution.

Methods

The analysis was carried out using a database obtained from a questionnaire administered to the patients with nasal symptoms containing personal data, clinical data and results of allergy testing (SPT). All variables found to be significantly different between patients with positive and negative SPT ($P < 0.05$), were selected for the logistic regression models and were analyzed with backward stepwise logistic regression. A second set of patients from another Institution was used to prove the model.

Results

The accuracy of the model in identifying, over the second set, both patients whose SPT will be positive and negative was high. The model detected 96 percent of patients with nasal symptoms and positive SPT, and classified 94 percent of those with negative SPT.

Conclusions

The data of the thesis have been preliminary to the creation of a software which could help the primary care doctors in diagnostic decision making process (need of allergy testing), in patients complaining of chronic nasal symptoms

Keywords:
Allergic rhinitis, Nonallergic rhinitis, Decision matrix, Logistic regression model, Receiver Operating Characteristic curve, Probability, Diagnostic decision making, Nasal symptoms, Skin prick test (SPT), Cognitive Errors

Introduction

The choice of a diagnostic path is one of the most important and intellectually challenging aspects of medical reasoning. When a general practitioner encounters a patient with nasal symptoms, faces a vast amount of information: the patient's lifelong personal, the medical history and the report of the nasal symptoms. In addition to this information, the general practitioner must have a good knowledge about allergic and non allergic rhinitis. However, it is often difficult, according to these data to make a path diagnosis.

Rhinitis is a very common disorder, affecting 20% to 40% of the western population. Rhinitis can be classified as being allergic (AR), infectious, and nonallergic/noninfectious (NAR) [1]. The exact figures are unknown, but most ENT and Allergist clinics report a 50-50 division between allergic and nonallergic patients in perennial symptoms [2]. It is difficult to differentiate allergic rhinitis (AR) from nonallergic rhinitis NAR). Such as bronchial asthma, rhinitis also has nonspecific hyperreactivity, and then nonspecific stimuli can induce nasal symptoms, in AR. On the other hand, specific stimuli can be indicate from the patients with NAR as the cause of nasal symptoms. For these reasons, the hyper-reactivity to specific or nonspecific stimuli, reported by patients with chronic nasal symptoms does not indicate to any form of rhinitis.

Chronic rhinitis, is a very prevalent disease that is often trivialized and therefore inaccurately diagnosed, leading to inadequate management and unnecessary health care expenditures.

When a general practitioner orders a allergy testing for a patient with chronic nasal symptoms, often did not take into consider the NAR. In fact, on the basis of his knowledge and experience, the AR is most frequent. Often, a patient history and physical examination will provide enough information to determine if a patient is allergic. However, only the presence of specific IgE, allows you to make a clinical diagnosis of allergy. Diagnostic tests for specific IgE, both *in vivo* and *in vitro*, should be used to diagnose and treat allergy only knowing the medical history of the patient. In other words, allergy testing should be performed only in patients with a real suspicion of allergy. There is broad consensus in the medical community, that this "remote practice of allergy" is unacceptable [3].

For all intents and purposes, the symptoms are a diagnostic test [4,5]. Therefore, on the one hand the result of allergy testing can be considered as the gold standard because reproducible and validated and the other side of the symptoms and clinical aspects of the clinical aspects reported by patients. In fact, the finding and the assessment of the clinical history and symptoms play a decisive role in shaping the decision of the general practitioner, to wish to run or not allergy testing [6,7].

The decision to perform allergy testing should not depend merely on generic impression. A number of critical methods are available to assess the significance of symptoms and of certain aspects of the clinical history. In addition, a critical evaluation is necessary to justify the execution of diagnostic procedures, including simple and relatively expensive test, as allergy testing in these days of limited resources, both for the health system and for patients [8,9].

Methods to evaluated diagnostic choices

A number of critical methods are available to evaluate diagnostic choices. In addition, critical evaluation is necessary so that use of given diagnostic choices can be justified in these days of limited resources both for patients and for public medical care.

The technical terms used for the four methods to be described are the decision matrix, logistic regression, the receiver operating characteristic (ROC) curve and information theory.

When a general practitioner evaluate the clinical characteristics of the patients, he has, on the basis of his knowledge and experience, a certain impression of its reliability.

Decision Matrix

By use of a decision matrix we can logically relate the results of a diagnostic test (e.g. clinical characteristics) to the clinica1 or pathologic outcome (e.g. result of Skin Prick Test (SPT). This type of analysis is most easily applied to the simple decision of whether SPT is positive, D+, or negative, D-, when the test (each variables of clinical characteristics) is present (i.e., positive), T+, or absent (i.e., negative), T - . When, as shown in Table 1, these two binary results are plotted on a two **X** two table to show the four possible combinations (indicated by a, b, c and d), a decision matrix is formed.

Table 1. Example of General Decision matrix.

Test results	Presence of disease		
	Present (D+)	Absent (D-)	totals
Present (T+)	a	c	a+c
Absent (T-)	b	d	b+d
Totals	a+b	c+d	

Each of the four combinations can be used to evaluate the test by comparing its results to the actual presence or absence of disease (i.e., four ratios may be formed). The so-called true-positive (TP) ratio is the proportion of positive tests in al1 patients that actually have the disease, or $\frac{a}{a+b}$ This value expresses probability (P) that patients with the disease will have abnormal test results, and can be written as the "conditional probability" P(T+ |D+) [*A "conditional probability" is written as a matter of convention, with a vertical bar before the given state or condition that is present or absent. It does not imply division*] - i.e., the probability that a patient with disease, D+, will have a positive test, T + . The true-positive ratio expresses the *sensitivity* of the examination. It measures the fraction of patients with disease that will be detected by the diagnostic test in question.

The false-positive (FP) ratio is the proportion of positive tests in all patients that do not have disease, or $\frac{c}{c+d}$. It is the probability that patients without disease will have abnormal test results, P(T+ |D -). The true-negative (TN) ratio is the proportion of negative tests in all patients that do not have the disease, or $\frac{d}{c+d}$. It is the probability that patients without disease will have negative test results, P(T-JD-). This ratio expresses the *specificity* of the examination. It measures the fraction of patients that will be correctly identified as having no disease. It is equal to (1 - FP ratio).

The false-negative (FN) ratio is the proportion of negative tests in all patients with disease, or $\frac{b}{a+b}$. I t is the probability that patients with disease will have negative test results, P(T - |D +). It is equal to (1 - TP ratio).

Obviously, a good diagnostic examination has a high TP ratio and a low FP ratio; it correctly identifies a large portion of diseased patients without incorrectly including patients without disease. The ratio of the TP ratio to the FP ratio is known as the likelihood ratio, L. Obviously, tests with high likelihood ratios are better discriminators of disease than those with low ones.

These test characteristics may be illustrated with a specific example of the results reported below.

In our study the SPT used for detecting the patients with specific sensitization Each clinical characteristic was examined considering the SPT. When the actual numbers as determined by the SPT examinations are put into the decision matrix, the following table emerges (Table 2).

Table 2. Correlation of sneezing and SPT

Sneezing results	Results		
	SPT+ (D+)	SPT+ (D-)	totals
Present (T+)	953	150	1103
Absent (T-)	8	248	256
Totals	961	398	

We may calculate the performances of each clinical characteristics as follows:

True positive ratio = P(T+|D+) = $\frac{953}{953+8}$ = 0.99

False positive ratio = P(T+|D-) = $\frac{150}{953+8}$ = 0.15

True negative ratio = P(T-|D-) = $\frac{248}{248+150} = 0.62$

False negative ratio = P(T+|D+) = $\frac{8}{8+953} = 0.0008$

Thus, the sneezing is 99 per cent *sensitive* and 62 per cent *specific*. It will detect 90 per cent of patients with SPT+ and will correctly classify 63 per cent of those with SPT- [10,11]

Logistic regression

Logistic regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). In our example, we have the clinical characteristics of each patient and on the other hand the result of the SPT. In other words, it is used to predict a binary response, dependent variable (i.e., the result of SPT) , from a binary or continue predictor (i.e. clinical characteristics of patient with chronic nasal symptoms). Logistic regression might be used to predict whether a patient has a given feature (e.g. SPT), based on observed clinical characteristics of the patient. In Table 2 we reported the variable examined. Logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is a noncase [12].

ROC curve

The single variable cannot be used to differentiate subjects with or without SPT+, because the clinical characteristics of the patients with SPT+ or SPT- are related with several variables as demonstrated by logistic regression. However, the Logistic Regression model was used to describe the relation between twelve explicative, continuous variables or dichotomic and a dichotomic dependent variable. Our main objective is to propose an unconventional working method to apply ROC curves to the determination of cutoff point on the continuous explicative variables, in situation in which the response variable have been analyzed by logistic regression [13].

Bayes's Theorem

Once a diagnostic test has been evaluated so that its characteristics (i.e., sensitivity and specificity) are known, it is possible to formulate new probability statements about the presence or absence of disease in a particular patient examined by the his clinical characteristics. These probability statements are called posterior or post-test probabilities because they reflect the test results. The clinical characteristic of a patient that has a clinical result (i.e. symptom) the probability of sensitization (disease) is written as P(D+|T+) [*the symbol "|" is read as given*]

and if he has a normal test result. it is written as P(D+|T-). Bayes's theorem is a technic that allows us to calculate these posterior probabilities that we wish to know from information that we already know beforehand ("a priori") about the implications of a diagnostic test [14]. For example, if we wish to estimate the probability of the SPT (disease) in a patient with a clinical characteristic (abnormal test) result we must know the probabilities that the diagnostic test will be positive in patients with and without disease - the TP and FP ratios and an estimate of the prior probabilities, P(D+) and P(D-). The following formula is used:

$$P(D+|T+) = \frac{P(T+|D+)P(D+)}{P(T+|D+)P(D+)+P(T+|D-)P(D-)} \quad (2)$$

Alternatively, if we wish to know the probability that a patient without that clinical characteristic (normal test) result has disease, we need to know the TN and FN ratios as well as P(D+) and P(D-). The relevant formula is:

$$P(D+|T-) = \frac{P(T-|D+)P(D+)}{P(T-|D+)P(D+)+P(T-|D-)P(D-)} \quad (3)$$

As a specific example illustrating these formulas consider the hypothetical test (Table 3, Fig. 1)

performed on a group of patients 30 per cent of whom are estimated to have disease. Let us assume that we have used point D as our cutoff point.

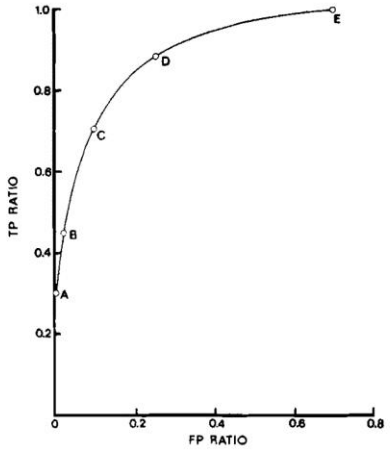


Figure 1. Hypothetical ROC curve

Table 3. Correlation of cutoff point and true-positive and false-positive ratios for a hypothetical Roc curve

Cutoff point	Proportion of patients having abnormal test results	
	Patients with disease (true-positive ratio)	Patients without disease (false-positive ratio)
A	0.30	0.00
B	0.45	0.02
C	0.70	0.10
D	0.90	0.25
E	1.00	0.70

The probability of disease in a patient with an abnormal test is calculated from equation (2) and is $P(D+|T+) = \frac{(0.90)(0.30)}{(0.90)(0.30)+(0.25)(0.70)} = 0.61$

Thus, the abnormal test has changed the probability of disease in a patient from 0.30 to 0.61, a factor of two. If, on the other hand, the patient has a normal test his probability of disease is calculated from equation (3) and becomes

$P(D+|T-) = \frac{(0.10)(0.30)}{(0.10)(0.30)+ (0.25)(0.70)} = 0.05$

A negative test has reduced the probability of disease from 0.30 to 0.05, a factor of six. In this context, the test is more useful in ruling out disease than in detecting it. The difference between posterior and prior probabilities is strongly dependent upon the true-positive and false-positive ratios for the diagnostic test. A nomogram relating both prior and posterior probabilities to these

ratios has been constructed for a wide range of test sensitivities. For tests that are "perfectly sensitive" (TP ratio = 1.0), a family of curves relating prior to posterior probability can be constructed for varying false-positive ratios [15].

Cognitive errors in diagnostic choice

In the last decade, the topic of medical error has stimulated a great deal of interest, not only among doctors and surgeons, but also among psychologists, economists and managers [16–19]. More recently, this topic has also gained the attention of those in the fields of Logic and the Philosophy of Science. Some epistemologists that have tackled the issue of error in the field of medicine have singled out two basic categories: that of errors committed in scientific research, when a researcher devises or accepts an unfounded hypothesis, and that of mistakes committed in the application of scientific knowledge, as is the case in clinical medicine, whereby doctors rely on knowledge held to be true at the time in order to understand an individual patient's signs and symptoms [20-22]. In this paper will treated deal exclusively with the latter, with particular reference to allergic rhinitis, that is to say the mistakes which general practitioner make while carrying out their day-to-day medical duties. For reasons of simplicity, in the text the terms 'error' and 'mistake' will be used as synonyms. Error can be considered to be any conclusion diagnostic or physiopathologic which deviates from what is held to be 'true', regardless of its consequences or only those events which lead to harmful consequences for the patient [23,24]. We analyze the errors that the general practitioner may commit in the course of his day-to-day activities when he confronts a problem with allergic.

Clinical errors in medicine have been largely attributed to the lack of, or flaws in observation. In other words general practitioner made mistakes because they failed to observe or pick up on certain signs, or because they misinterpreted the results of diagnostic tests [25]. This opinion is based on the conviction that the most important elements, in clinical methodology, consist in the objective and thorough reporting of the patient's condition and his/her personal and clinical history. However, the problem of clinical error is more complex and the key factor involved in the errors of general practitioners is not always report the facts. Medicine is not a science, but it is a scientific discipline and, like all scientific disciplines, consists of two parts that are closely intertwined: an empirical part, where the real-world events are observed and described, and a rational part, in which the various phenomena are placed in relation. Therefore, the clinical errors can derive both from mistakes in the recording of empirical phenomena and from errors in the physician's thought processes [26,27]. Today the Journals of Allergology gave little attention to errors in general practitioner' reasoning during and after the examination of the patient. This has led to the recurrence of mistakes are often due to incorrect reasoning rather than an inadequate observation of their patients.

A study has examined about 100 errors in internal medicine, 28 cases were due to cognitive errors exclusively and in 72 cases cognitive errors also played a part. The study has

demonstrated that the majority of these mistakes were due to problems in the elaboration of the information available, in other words in the process of reasoning based on these facts [28].
We will demonstrate that the main duty of the general practitioner should be to reason well, that is to say to reason in a correct way, in the interest of the patient. It goes without saying that reasoning is a process by which one arrives at a conclusion from an initial premise, after having elaborated certain hypotheses; not all reasonings however have the same value and from this point of view one can distinguish (1) demonstrative reasoning, (2) argumentative thought, (3) deceptive reasoning or fallacies [29,30].
The objectives methodological in clinical practice are:

---

(1) to classify patient's disease

(2) to understand the pathological phenomena in patient

(3) to predict the progression of the diseases of patient

(4) to modify the predicted progression using medicines, diet, specific therapy, when it is indicated

---

The errors related to the first three objectives are cognitive errors, that is to say errors which are linked to the knowledge of general practitioner, and which therefore concern the question of truth. Errors related to the fourth objective, on the other hand, are operational errors, perhaps most easily attributable to the Allergist. Indeed, it appears obvious that a diagnostic hypothesis or explanation can be either true or false, whereas a treatment cannot be false. It can be inadequate, insufficient, ineffective, useless, damaging, dangerous, outmoded, misdirected and so on. The errors committed by general practitioner can be classified in various ways and each classification has its advantages and disadvantages [28, 31-33]. A reasonable classification of the errors committed by general practitioner has been reported below.

Classification of errors of the general practitioner

Cognitive errors:
Errors concerning medical knowledge
Errors concerning clinical methodology
Errors of reasoning
Operative errors:
Errors in the analysis of symptoms
Decision making errors
Errors in the therapeutical approach*
** Allergist's error*

We will analyze only the errors of reasoning, it should be remembered that, following the widely accepted methodological tradition, diagnostic argumentations are based principally on calculations of probability which are in turn based on the famous theorem set out in 1763 by the Reverend Thomas Bayes [14]

Bayes' Theorem in form applied to clinical medicine

$$P\,(M|S) = P\,(M|S)\,\frac{P(M)}{P(S)}$$

In medicine, this theorem allows us to go from calculating the probability of finding certain signs in the presence of a certain illness to calculating the inverse probability, that is to say the probability of finding a certain illness in the presence of certain signs. Indeed, after having gathered the necessary information regarding signs and symptoms, each of which has a specific probability of being present in various different illnesses (probability of the pathologists) and based on knowledge of the prevalence of those particular illnesses and those particular signs in a given population (probability of the epidemiologist), this theorem allows us to calculate the probability that an individual patient displaying those particular signs might be suffering from a specific illness (probability of the clinicians). In fact, general practitioners often make

mistakes because they do not know or misapply Bayes' theorem. Either they misjudge the sensitivity, specificity or the predictive value of a specific sign, or they do not take into account or miscalculate the prevalence of a specific illness among a specific population, or they believe that probability of the pathologists is the most important, or finally because they do not take certain pathologies into account.

We will make an example of error in Bayesian reasoning. A general practitioner examines a woman of 82 years that complains of chronic nasal symptoms. The doctor said (based on his experience and the literature) assesses that the probability of the AR for a woman of her age, habits and family background is 1%, whereas the probability of it being NAR is 99%. The doctor sends the patients for the assay of specific IgE (Phadiotop) and the medical laboratory found the Phadiatop positive. In terms of the diagnosis, our doctor's problem is therefore to calculate the probability of his patient actually suffering from AR. He makes inquiries about the accuracy of the specific IgE and discovers that this investigation correctly shows up a sensitization in 69.7% of confirmed allergy cases and that NAR are correctly diagnosed in 95% of cases. Using Bayes' theorem, our doctor applies the following formula to his calculations:

$$\text{P (AR|Phadiatop pos)} = \frac{P(Spos|RA)*P(RA)}{[P(Spos|RA)*P(RA)]+[P(Spos|NAR)*P(NAR)}$$

and concludes that the probability of the patient having AR is approximately 8% [34].

Many doctors think that the probability of AR would be in one of their patients under these same conditions about of 70%. The reasoning is that the probability that the patient has an AR is equal to the probability of obtaining a positive result in confirmed cases of AR. This is the probability of the pathologist, but not that of the clinicians or, in other words retrospective accuracy with predictive accuracy.

The importance of the clinical characteristics was used to assess clinical analyze the probability of a diagnosis rather than the choice of a diagnostic path. When two different diagnosis for a patient seem to have the same potential value, the decision of the physician is often described as a "toss-up", and more often than a physician decide to follow the one that seems most obvious. But that's not always the most obvious decision coincides with the right. Many of these decisions are related to inductive reasoning. [20,23,30].

Perhaps the most common of which is the so-called 'fallacy of statistical bias'. This consists in making inductive generalizations based on a sample which is known to be unrepresentative of the population, or a sample which cannot objectively be considered representative. In clinical practice

this type of mistake happens when, for example, after having observed six subjects that have chronic nasal symptoms and show a worsening of nasal symptoms in the spring, and these patients all have allergic rhinitis, one concludes that the worsening of symptoms chronic nasal in the spring is to be associated with allergic rhinitis. A similar, yet more serious mistake is the fallacy of causal correlation, or rather the fallacy known as ‘‘post hoc, ergo propter hoc’’. In this case, a causal role is attributed to a certain event (event A) simply because it had preceded another event (event B) which is then seen as being influenced by the first event.
For example

This patient in the spring , on a windy day, went to the countryside

After some hours the patients hours the following nasal symptoms: rhinorrhea and nasal obstruction.

The pollens in the air were the *cause* of the nasal symptoms.

Another mistake of logic consists in the confusing of cause and effect.
For example:

The patient is suffering from a significant inattention and he have a diagnosis of allergic rhinitis.

The patients takes antihistamines.

It is well known fact that some antihistamines, but not all, can cause a significant inattention.

The significant inattention is caused by the antihistamines.

The argumentation set out here is not at all conclusive, one would be equally justified in maintaining that the patient suffers from significant inattention caused by the persistent nasal obstructions of allergic rhinitis.

A frequently committed error in complex cases is the socalled “Petitio principia”, whereby one assumes to be true that which one wishes to demonstrate.
For example:

The chronic nasal symptoms of the patient were caused by an allergic reaction

How do you know that the nasal chronic symptoms are brought about by an allergic reaction?

Can’t you see that those suffering from AR, have these symptoms?

Indeed an important fallacy in clinical practice is the so-called fallacy of division, which can be observed when, in order to formulate a diagnosis, the doctor relies on differential diagnosis. As we know, differential diagnosis is based on the disjunctive syllogism. The reasoning usually used in differential diagnosis can take one of the two following forms:

I

The patient P is suffering from illness M1 or illness M2 or illness M3

The patient is not suffering from illness M2 nor illness M3

—

The patient P is suffering from illness M1

II

The patient P is suffering from illness M1 or illness M2 or illness M3

The patient P is suffering from illness M1

—

The patient is neither suffering from illness M2 nor illness M3

Both of these processes are subject to the fallacy of disjunction when the number of illnesses from which the patient might be suffering is greater than the number of illness effectively taken into account. In fact, if one affirms that:

P's chronic nasal symptoms could be due to a allergic reaction or a non allergic reaction

---

and one does entertain the possibility that these chronic nasal symptoms might also be due to a viral rinitis or an anatomical variants in the ostiomeatal complex then the demonstration of an allergic reaction or the demonstration of the absence of a non allergic reaction could lead to an incorrect conclusion. The reasoning II is based on an inconclusive (weak) disjunction, the premiss simply affirms that at least one of the possibilities put forward is true. However, the patient might be suffering from illness M3 and illness M1, or by illness M3 and by illness M2 [35].

Many errors are due to fallacies of deductive logic, logic which the doctors use on a day-to-day basis examining patients in order to envisage the consequences of the various diagnostic or physiopathologic hypotheses. To illustrate these errors, errors which are often committed unwittingly, we will consider some simple deductive arguments which are partly valid and partly invalid:

Deduction I

All individuals with shock have low blood pressure (t)

All those with anaphylactic shock have shock (t)

All those with anaphylactic shock have low blood pressure (t)

Clearly both of the premises in this deduction are true, the conclusion is true and the deductive argumentation is valid.

Let us now turn our attention to another deductive inference:

Deduction II

All individuals with low blood pressure have tachycardia (f)

All those with tachycardia have low blood pressure (f)

All those with anaphylactic shock have tachycardia (t)

In this second inference, the premises are false, the conclusion is true but the argumentation is identical to the previous one and the deduction is, therefore, perfectly valid.
Now let us consider a third deductive inference:

Deduction III

All individuals with low blood pressure have tachycardia (f)

All those with low blood pressure have tachycardia (f)

All those with anaphylactic shock have tachycardia (t)

Also in this case the premises are false and the conclusion is true. The conclusion however, despite being true, is not guaranteed by the premise and so, therefore, the deduction is not valid.
Now let us ponder a fourth inference:

Deduction IV

Some individuals with low blood pressure have tachycardia (t)

Some individuals with anaphylaxis have low blood pressure (t)

Some individuals with anaphylaxis have low blood pressure and have tachycardia (t)

In this case the premises and conclusion are both true, but the conclusion is not guaranteed by the premises and so, therefore, the deduction is not valid.
And so turning to a fifth deduction:

Deduction V

Some individuals with anaphylaxis have low blood pressure (t)

All individuals with anaphylactic shock have low blood pressure (t)

Some individuals with low blood pressure have anaphylaxis (t)

In this case the premises and the conclusion are both true and the deduction is valid.
Some fundamental conclusions can be drawn from these examples of deductive reasoning:

(a) it is an error of logic to infer the truth of the premise from the truth of the conclusion.

(b) it is an error of logic to infer from the truth of the conclusion the validity of an argument.

(c) it is an error of logic to infer from the falsity of the conclusion the invalidity of an argument.

(d) it is an error of logic to infer from the falsity of the conclusion the falsity of the premise.

Applying these general conclusions to clinical logic, it can be affirmed that:

(a) based on a true diagnosis it cannot be concluded that the initial premise were true

---

(b) based on the truth of the diagnosis, it cannot be concluded the reasoning was correct

(c) based on an incorrect diagnosis, it cannot be concluded that the reasoning was incorrect

(d) based on an incorrect diagnosis, it cannot be concluded that the initial observations were false

---

The deductive argumentations which we have been looking at until now concerned classes of individuals and were aimed at affirming that a specific characteristic can be attributed to a given class, or at least to a part of the individuals that make up that class: for example the characteristic of low blood pressure to those suffering from tachycardia or the characteristic of tachycardia in those suffering from shock.

There are, however, other deductive argumentations which do not regard the attribution of characteristics, but which aim to confirm or refute a given hypothesis. Some of these deductions are clearly fallacies. Among these, probably the most common deductive fallacy is the fallacy of confirmation, or the fallacy of the affirmation of the consequent. Here is an example:

If this patient is allergic, he will present specific IgE

This patient present specific IgE

This patient is therefore allergic.

Clearly this argument is not valid because the patient with specific IgE might not be allergic The detection of these antibodies by *in vivo* or *in vitro* methods, nevertheless, merely indicates the existence of sensitization to given allergens; their presence does not always coincide with clinically significant allergic disease. Thus the diagnosis of allergic disease is usually based on a combination of information obtained from the medical history, which by itself cannot be considered as the *gold standard*, with that obtained from *in vivo* and/or *in vitro* tests [36].

This is illustrated in the following example:

If nasal symptoms patient's disappear after the administration of antihistamines, the nasal symptoms are therefore due to an allergic reaction.

The nasal symptoms does not disappear after the antihistamines.

Therefore the patient is not suffering from an AR.

As has been demonstrated, errors of reasoning can take many forms. We can err in many different ways and often we do not realize that we are doing so, we take for conclusive argumentations only assertions a weak support for our argumentations or our hypotheses. Herein lies the sneaky way with which errors creep into our minds: we are convinced that we have finally reached a diagnostic truth or a true physiopathologic explanation, whereas in fact we have simply been chasing a shadow. Until now, we have considered the errors of reasoning committed in clinical medicine. Now it is necessary to point out a different type of errors that are part of the psychology of thought. In recent years, in fact, a new field of research has been developed to analyzing the way in which doctors, economists, and other carry out reasoning tasks [37-39]. This research has shown that in practical reasoning tasks humans commit many logical errors and that the participants' responses are greatly influenced by the problem context and its content, despite the logical irrelevance of these aspects. Studies by Tversky and Kahneman have led to the conclusion that human rationality is limited by cognitive conditioning. In a situation of uncertainty, this cognitive conditioning leads the decision maker into using simplification strategies - known as heuristic choice - which are linked to systematic errors of judgement. Even the experts are not immune to this tendency and in various applied situations move away, to a greater or lesser extend from the formal rules of logic [40,41].

These errors in applied tasks of reasoning are profoundly different from those we have considered until now; while the aforementioned represent errors of logic, errors that is which derive from the failure to observe the rules of thought, the mistakes which we are about to discuss depend on the way in which the problem is presented (framing). In order to illustrate this we will look at just two examples. When a physician has to formulate a diagnosis, he puts forward a certain number of hypotheses and assigns a certain probability to each one of these. For example, faced with a case of chronic nasal symptoms he conjectures that his patient might be suffering from allergic rhinitis or non allergic rhinitis or sinusitis or an anatomical variants

in the ostiomeatal complex. A group of doctors were asked to assign a probability to each of these hypotheses. Their responses were as follows: allergic rhinitis was given a 55% probability, non allergic rhinitis a 10% probability, sinusitis 30% probability, and finally anatomical variants in the ostiomeatal complex a 5% probability. As required by the theory of probability, the sum of these figures is 1. The same group of students, faced with an identical case of nasal chronic symptoms, were given a similar task of assessing the probability for each one of the illnesses in same set as before. On this occasion however, allergic rhinitis was divided into 3 subgroups: seasonal rhinitis, perennial rhinitis and perennial rhinitis with seasonal exacerbation. The second part of the experiment brought some unexpected results, by dividing the general heading of allergic rhinitis into 3 subgroups, the sum of the probabilities assigned to these subgroups was greater than the probability assigned to the general heading under which they were included, whereas the probability of the remaining hypotheses (non allergic rhinitis or sinusitis or an anatomical variants in the ostiomeatal complex) remained unchanged. As such the total probability of all the hypotheses combined paradoxically became greater than 1. This experiment shows that when the spotlight is turned onto a specific diagnostic hypothesis, one which had previously been overshadowed under a wider umbrella hypothesis, and a specific probability is assigned to this hypothesis, students often fail to review the probability of the various remaining diagnostic hypotheses. This inevitably leads to miscalculations [42].

Let us look at another experiment. The following situation is described to a group of students of Medicine: a doctor is called upon mid-flight to examine a 60-year-old passenger that presented a generalized urticaria, after eating peanuts. The patient has hypertension treated with beta-blockers [43]. The First Aid kit contains a sphygmomanometer, which gives a SBP reading of 100 mmHg. The group of students are asked whether they would recommend the administration of adrenaline or whether was sufficient administration of antihistamines. The majority of students (89%) reply that it was necessary to administer adrenaline. In the second part of the experiment the students are confronted with the same scenario, except this time the First Aid kit does not contain a sphygmomanometer.

The cabin crew are sure that they have seen it and insist on conducting a search. This extensive search leads to the discovery of the apparatus. The SBP is finally measured and a reading of 100 mmHg is obtained. The students are asked the same question as in the first scenario. Their response in this case differs from the previous one, even though both situations are essentially the same. In the second case a greater number of doctors (85%) recommended the administration of antihistamines. This second experiment shows that the importance given to a

clinical sign varies depending on whether it comes to light immediately or whether it is discovered thanks to a deliberate and determined search [44].

The results of these and many other psychological experiments throws new light on the rational performance. It has always been maintained that in their daily lives people's beliefs, and also those of doctors, are clear cut and unchanging over time. It has also been maintained that the acquisition of new information could but improve our suppositions about reality and our ability to arrive at an affirmation of the truth. In reality, psychological studies based on practical reasoning tasks show that the situation is really not that simple: the beliefs are not impartial with regards to their actions, these beliefs are not always clear cut and are not constant over the course of the investigation. On the contrary, often these beliefs are actually formed during the investigative and decision making process. As has been stated ''individuals' priorities are subject to change and a small difference in circumstances can sometimes alter people's preferences and led them to make alternative decisions''.

All things considered, we can conclude that general practitioner often make mistakes because, unknowingly, they fail to reason correctly. In the past, it was thought that the errors committed in medicine were essentially due to defective or incomplete observation and that once general practitioner had learnt to 'observe well' they would be safe from committing errors.

Unfortunately, this is not the case. Even after an general practitioner has carried out a thorough and correct examination of his patient, he can still make a mistake. These mistakes can occur in two basic ways: either because he does not observe the laws of formal logic and so falls into the trap of one of the many fallacies which logic shows us how to avoid, or because his practical rationality does not match theoretical rationality and so his reasoning becomes influenced by the circumstances in which he finds himself.

In conclusion Medicine, unlike Science is fallible because it is human. This is beyond doubt, yet it should be added that clinical medicine is doubly fallible: both because it is the work of man and because these men must intervene to resolve the problems of others men in emotionally demanding circumstances [45].

Experimental data

Chronic rhinitis is typically classified as allergic rhinitis (AR) if the symptoms and triggers correlate with a specific IgE-mediated response, or as non-allergic rhinitis (NAR) if symptoms are induced by irritant triggers in the absence of specific IgE-mediated responses [46, 47]. AR is a common condition affecting 5 − 40% of the general population and there is evidence that its prevalence is increasing [1]. Rhinitis is an inflammation of the nasal membrane that causes periods of nasal discharge, sneezing, and congestion that persist for at least two hours per day [1, 46,47]. Rhinitis is considered allergic when allergen-specific IgE initiate the immunologic reaction that causes symptoms, while it is non-allergic if allergen-specific IgE are negative [48]. Diagnostic allergy tests attempt to detect specific IgE, which cause the nasal symptoms, binding common allergens, such as house dust mites, pollens, animal proteins, and mold spores [49]. Primary care doctors are usually the first to encounter patients with chronic nasal symptoms, but they are often uncertain about how to differentiate between allergic and non-allergic forms of the disease. They normally require a consultation with an allergy specialist if nasal symptoms have been present for more
than two years, and they occur cyclically [50].

The availability of a short questionnaire for the diagnostic decision, that correlates with the positive or negative allergy test, may serve to modify and rationalize the current approach taken by primary care doctors to evaluate patients with chronic nasal symptoms. In other words, we will try to answer the question: is it necessary for patients to undergo allergy testing? This decision will be made considering their demographic and clinical characteristics. Due to its simplicity, high sensitivity, rapid interpretation, and a relatively low cost, skin prick test (SPT) was until recently, often recommended by primary care doctors. But because of the current state of the economy and health care system problems, health care expenditure has fallen. In the Italian health care system, the total cost of an allergy test and *in vivo* testing is €44, considering the cost of the allergy extract and the allergist's charge (allergist's time for the medical history, clinical examination, and SPT [51]). This is paid by the patient, that possibly does not need an allergy test.

An important contribution to the rhinitis diagnostic decision process can be provided by the examination of a database performed on a wide sample of patients with chronic nasal symptoms. The crucial point is how to examine the data obtained from the database in order to assemble a questionnaire that will facilitate the diagnostic decision process of primary care doctors for new patients with chronic nasal symptoms.

A considerable amount of scientific production is directed at the exploitation of databases or questionnaires in order to implement models and algorithms useful to the assessment, assistance in medical diagnosis, and treatment of allergic rhinitis and respiratory diseases. In the early 80s Pantin and Merrett (1982) [52] applied a computer system to predict the IgE-mediated allergies of patients that answer a standard allergy questionnaire by referring to a database compiled from previous patients' answers and their IgE antibody profiles. Subsequently, other computer systems have been published [53-56]. However, none has been used in the clinical practice.

Our objective was to evaluate accuracy of SPT results, by the analysis, through a logistic regression model, of a database of 1359 patients with chronic nasal symptoms. The performance of the model was validated through a data set obtained from another medical institution of the center of Italy. Between south (Palermo, Sicily) and center (Rome, Lazio) of Italy there are very different aero-biological reality, and then different sensitivities in allergic patients [57, 58]. The results of this model could be useful for primary care doctors to start the correct diagnostic process, without interfering with the role of the doctor that, in our opinion, cannot be replaced by a computer.

Patients and methods

Description of the database

The original database consists of 1511 patients, consecutively seen and evaluated in the outpatient allergy office of the 'Dipartimento BioMedico di Medicina Interna e Specialistica' (Di.Bi.M.I.S.) (ex Dipartimento di Medicina Clinica e delle Patologie Emergenti) of the University of Palermo, Italy. The database were previously used to analyze the characteristics of allergic rhinitis disease [4]. Of the 1511 patients with nasal symptoms reported in the previous study, we performed the current analysis evaluating 1359 patients that had completed the diagnostic process. The data were obtained from a questionnaire administered to the patients and containing their personal information, the main features about nasal symptoms, and the results of allergy testing. All questionnaires were administered by two Allergists (MSLB and SLP), who, if necessary, clarified to patients the meaning of the questions. All patients performed *in vivo* allergy testing: skin prick test (SPT), using a standard panel of allergens [4, 14]. Results of the SPT were confirmed by the assay of serum specific IgE. We analyzed the data of 25 different input variables for each patient of which: 3/25 were continuous variables, 4/25 were ordinal variables and finally 18/25 were dichotomous variables.

Statistical analysis

We compared each input variable between SPT positive and SPT negative output variable, using the student t test or Mann-Whitney test for continuous variables, depending on the distribution of the data, the χ^2 test for the dichotomous variables, and finally the Mann-Whitney test for the ordinal variables [59].

All variables found to be significantly different between patients with positive SPT and negative SPT ($p < 0.05$) were selected for the logistic regression model and analyzed with backward stepwise logistic regression. The goodness of fit of the logistic models was assessed using the Hosmer and Lemeshow test [60]. Several multiple logistic regression models were tested in order to determine the most significant and simplest model with the best available fit for the ~~data~~ available data. Based on the regression coefficients obtained for each significant factor chosen by the logistic regression, a predictive probability equation was used to generate predicted probability of positive SPT for each individual. The probability of having a positive SPT in patients with nasal symptoms (p) was then calculated using the following equation:

$$p = \frac{e^y}{1+e^y}$$

where

$y = \text{constant} + x_1\text{variable}_1 + x_2\text{variable}_2 + x_3\text{variable}_3 + \ldots$

The area under curve (AUC) of the receiver operating characteristic (ROC) [61], has been evaluated applying the equations indicated by Fletcher and Fletcher [18]. We then assessed the performance of the algorithm by calculating its sensitivity (Se), specificity (Sp), predictive positive value (PPV), predictive negative value (NPV), positive likelihood ratio (LR+), negative likelihood ratio (LR−), accuracy (A), geometric mean (GM) of the identified true rates of positive SPT (Se) and negative SPT (Sp), and diagnostic odds ratio (DOR).

Validation of the logistic regression model

To evaluate the probability of the previous obtained logistic model to identify correctly patients with positive or negative SPT [62], we analyzed the data of 88 adult patients with chronic nasal symptoms, consecutively seen and evaluated in the outpatient allergy office of the Dipartimento di Pediatria Ospedaliera G.B. Grassi of Rome, Italy. The data consisted of patients' personal information, main features of nasal symptoms and result of SPT. The evaluation has been performed in blind, then the physicians of Palermo did not know the result of the SPT performed in Rome and the physician in Rome did not know the result obtained from the responses to questionnaire, examined with the algorithm obtained from logistic model of the database of Palermo.

Ethic Committee

The study was approved by the Institutional Review Board of the Dipartimento di Medicina Interna e Specialistica [now Dipartimento BioMedico di di Medicina Interna e Specialistica (Di.Bi.M.I.S.)] of the University of Palermo, Italy, and it was conducted according to the Declaration of Helsinki. Authorization of the study was not required according to our institutional policy and the ethical committee of our institution, as procedures done were part of routine diagnostic testing. However, written informed consent for the study was obtained from each patient in compliance with our institutional policy.

Results

Statistical analysis of database of Palermo

On the basis of the analysis reported in the previous section we selected the variables used to make the logistic regression model. Seventy-one percent of the subjects, present in Palermo database, were positive on SPT (n = 961). The clinical characteristics of the study population divided into positive and negative SPT subgroups are summarized in Table 4.

Table 4. Demographic and Clinical characteristic of the population

Variable	SPT positive (961)	SPT negative (398)	*P*
Age*	32.83 (11.50)	39.88 (15.06)	< 0.0001
Sex†	488 (50.78)	131 (32.91)	< 0.0001
Years of nasal symptoms*	7.11 (6.63)	7.5 (8.09)	0.35
Age of onset nasal symptoms*	25.81 (11.42)	32.52 (15.69)	< 0.0001
Family history†	336 (34.96)	103 (25.90)	0.0011
Smoker†	187 (19.46)	78 (19.60)	0.95
Ex smokers†	82 (8.53)	36 (9.05)	0.76
Passive smoker†	216 (22.48)	102 (25.63)	0.21
Trigger of nasal symptoms†	929 (96.67)	296 (74.37)	< 0.0001
Presence of pet†	109 (11.34)	6 (1.51)	< 0.0001
Nasal symptoms outdoor†	442 (45.99)	74 (18.58)	< 0.0001
Perennial nasal symptoms†	802 (83.45)	328 (82.41)	0.64
Sneezing†	953 (99.16)	150 (37.69)	< 0.0001
Severity of sneezing‡	3 (2–3)	0 (0–1)	< 0.0001
Rhinorrhea†	891 (92.72)	372 (93.47)	0.62
Severity of rhinorrhea‡	2 (1–3)	2 (1–2	< 0.0001
Nasal itching†	910 (94.70)	182 (45.73)	< 0.0001
Severity of nasal itching‡	2 (1–3)	0 (0–1)	< 0.0001
Nasal obstruction†	856 (89.07)	386 (96.98)	< 0.0001
Severity of nasal obstruction‡	2 (1–3)	2 (2–3)	< 0.0001
Conjunctivitis†	356 (37.04)	50 (12.56)	< 0.0001
Duration of nasal symptoms†	370 (38.50)	173 (43.47)	0.08
Severity of nasal symptoms†	644 (67.01)	129 (32.41)	< 0.0001
Clinical response to antihistamines†	499 (51.93)	43 (10.80)	< 0.0001

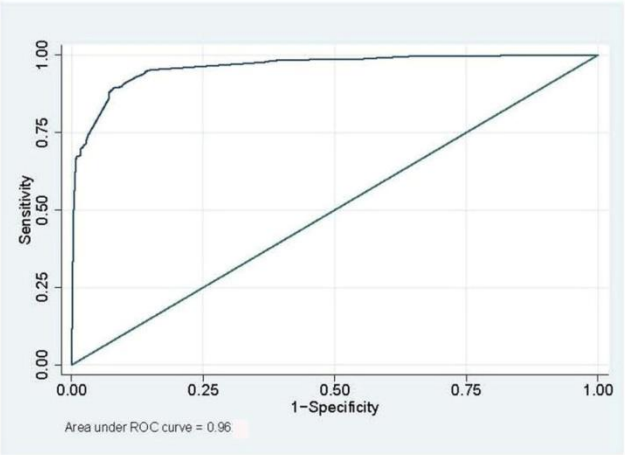
*means (standard deviations); confidence level 99%, equal variances hypothesis
† frequencies of the value 1 (relative percentages)
‡ medians (25^{th}–75^{th} percentiles)

The best logistic model included the following variables: age, sex, presence of family history of IgE-mediated allergic disease, if nasal symptoms were related to allergens, if the subject is in contact with pets, where the nasal symptoms occur, presence of sneezing, presence of nasal itching, presence of nasal obstruction, presence of conjunctivitis, severity of nasal symptoms, and presence of a clinical response to antihistamines. Table 5 shows the results of logistic regression retaining only the significant covariates in this model. The presence of a clinical response to antihistamines , severity of nasal symptoms, in contact with pet, nasal symptoms related to allergens, presence of nasal itching, presence of sneezing, presence of conjunctivitis, nasal symptoms occur outdoor, and male (sex) were associated with positive SPT, while older age, and presence of nasal obstruction, were associated with negative SPT.

Table 5. Variables that significantly distinguish subjects with SPT^+ from subjects with SPT^- by backward stepwise logistic regression

Variable	Odds ratios	95%CI	*P*
Age	0.96	0.94 to 0.96	< 0.0001
Sex	1.70	1.13 to 2.56	0.01
Family history	1.71	1.10 to 2.64	0.02
Trigger of nasal symptoms	6.08	3.22 to 11.49	< 0.0001
Presence of pet	7.24	2.40 to 21.81	0.0004
Nasal symptoms outdoor	1.96	1.23 to 3.14	0.005
Sneezing	1.13	2.66 to 6.43	< 0.0001
Nasal itching	5.20	3.32 to 8.14	< 0.0001
Nasal obstruction	0.19	0.07 to 0.51	
Conjunctivitis	3.88	2.32 to 6.48	< 0.0001
Severity of nasal symptoms	18.82	10.98 to 32.28	< 0.0001
Clinical response to antihistamines	10.37	21.68 to 75.17	< 0.0001

We performed ROC analysis for all variables that were found to be independently associated with SPT positive and SPT negative in order to identify the optimal cutoff point useful in predicting SPT positive. For the optimal cutoff point useful in predicting SPT positive, based on the above-mentioned logistic regression model, the following equation was generated and used in order to calculate predicted probability (pi) of having patient SPT positive: $\left(\frac{p\mathrm{i}}{1-p\mathrm{i}}\right)$ = (−3.32) +[− 0.05 (age)] + [0.53 (sex)] + [0.54 (presence of family history of IgE-mediated allergic disease)] + [1.80 (if nasal symptoms were related to allergens)] + [1.98 (if the subject is in contact with pets)] + [0.67 (nasal symptoms occur outdoor)] + [1.42 (presence of sneezing)] + [1.65 (presence of nasal itching)] + [− 1.64 (presence of nasal obstruction)] + [(1.36 (presence of conjunctivitis)] + [2.94 (severity of nasal symptoms)] + [3.70 (clinical response to antihistamines)]. This algorithm resulted in high predictive accuracy, considering as the best cutoff > 0.70: sensitivity 0.90 (95% CI [0.87, 0.92]), specificity 0.88 (95% CI [0.83, 0.92]), positive likelihood ratio 7.33 (95% CI [4.70, 10.30]), negative likelihood ratio 0.11 (95% CI [0.09, 0.14]) and area under curve 0.96 (95% CI [0.94, 0.97]) with a significance level < 0.0001. We reported in Figure 2 the performance of ROC curve of the model based on SPT results present in database of Palermo and the patients classified with the algorithm [17, 18]. Our results confirmed that the model can be considered as a diagnostic decision making tool.



Metric values of logistic regression model

Metric	Value	95% CI
Sensitivity	0.90	0.87 to 0.92
Specificity	0.87	0.83 to 0.91
Prevalence	0.70	0.67 to 0.73
Positive predictive value	0.94	0.92 to 0.96
Negative Predictive value	0.78	0.73 to 0.83
Positive likelihood ratios	7.33	4.70 to 10.29
Negative likelihood ratios	0.11	0.08 to 0.14
Accuracy	0.89	0.87 to 0.91
Geometric mean of the individuated *Se* and *Sp*	0.89	0.86 to 0.91
Diagnostic odds ratio	65.69	34.5 to 100.91

Validation of the logistic regression model

We analyzed a new database obtained from 88 adult patients with chronic nasal symptoms, consecutively seen and evaluated clinically and with SPT, in the outpatient allergy office of the Dipartimento di Pediatria Ospedaliera (Pediatric Department) G.B. Grassi of Rome, Italy. We used the input variables of logistic model, obtained from the Palermo database. In this way we evaluated how many patients with chronic nasal symptoms in Rome would be classified correctly by the logistic model, comparing the results with the results of SPT performed on these patients. We, then, evaluated the probability of the logistic model to identify correctly patients with positive or negative SPT. The results of SPT performed on these patients and the predicted SPT with the logistic model are reported in Table 6.

Table 6. Contingency table of the results of the logistic regression model, compared to allergy testing in patients of Rome and metric values.

From logistic model	From allergy testing *in vivo*		
	Positive SPT	Negative SPT	Total
Predicted positive SPT	53	2	55
Predicted negative SPT	2	31	33
Total	55	33	88

Metric	Value	95% CI
Sensitivity	0.96	0.87 to 0.99
Specificity	0.93	0.80 to 0.99
Prevalence	0.62	0.51 to 0.72
Positive predictive value	0.96	0.87 to 0.99
Negative Predictive value	0.93	0.79 to 0.99
Positive likelihood ratios	15.90	4.91 to 57.17
Negative likelihood ratios	0.03	0.01 to 0.13
Accuracy	0.95	0.91 to 0.99
Geometric mean of the individuated *Se* and *Sp*	0.95	0.83 to 0.99
Diagnostic odds ratio	410.75	45.27 to 4801.20

The logistic regression model is 96 percent sensitive, and 94 percent specific. It will detect 96 percent of patients with nasal symptoms and positive SPT and classify 94 percent of patients with negative SPT (Table 7).

Table 7. Positive and negative probabilities of the predictive logistic regression model

	Value	95%CI
P^{+}_{post}	0.96	0.84 to 0.99
P^{-}_{post}	0.94	0.74 to 0.99

Conclusions

The model performed well in predicting the result of SPT in individuals with chronic nasal symptoms (ROC curve areas is 0.95 for our logistic model). The logistic regression model had all the metric values greater than 80%, at ideal thresholds. The logistic model is simple and easy to interpret. Logistic regression yields a regression equation with coefficients for each significantly associated covariate. This is the first study that considering the clinical features of patients with nasal chronic symptoms predicted the result of SPT. Using our model, the primary care doctors could have a simple tool which might help, in the selection of patients that will benefit from the performance of allergy testing. We do not want to take the doctor's place but we would like to provide a tool that can help to make a decision. Our model correctly classified 96 percent of patients that will have a positive SPT, and 94 percent of those that will have a negative SPT. The use of our model would rationalize the access of patients with chronic nasal

symptoms, to allergy clinics in the public health system. This consideration would reduce the direct cost of the patients and the indirect cost of public health system.

Prior study examined patients with allergic rhinitis and non allergic rhinitis, however these included a smaller number of variables than our analysis and may have excluded important input variables [53-56]. To our knowledge this is the first study for prediction of the results of SPT as positive or negative on the base of the clinical history. We believe this approach produces the model that could be the most useful for primary care doctors, without interfering with the role of the doctor that, in our opinion, cannot be replaced by a computer. A part of the results of this thesis was recently published [63].

We translate our logistic model into a software (ARSTAT©), that should guide the general practitioner in the choice of the diagnostic reasoning for the patient with chronic nasal symptoms.

Acknowledgements

This study was supported by grants from MIUR (Italian University and Research Ministry) (former 60% funds) to Gabriele Di Lorenzo No support was received from the pharmaceutical and diagnostic industry. The authors declare they have no competing of interest. I would like to thank Professor Peter Dawson for the accurate proof-reading and Ing. Valerio Lacagnina of the Dipartimento di Scienze Statistiche e Matematiche “Silvio Vianelli”, Università degli Studi di Palermo, Italy for the software ARSTAT©.

References

1. Bousquet J, Khaltaev N, Cruz AA, Denburg J, Fokkens WJ, et al. Allergic Rhinitis and its mpact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). Allergy. 2008; 63 (Suppl 86) :8-160.
2. Fokkens WJ. Thoughts on the pathophysiology of nonallergic Rhinitis. Current Allergy and Asthma Reports 2002, 2:203–209.
3. Poon AW, Goodman CS, Rubin RJ. In vitro and skin testing for allergy: comparable clinical utility and costs. Am J Manag Care. 1998; 4: 969-85.
4. Li JT, Andrist D, Bamlet WR, Wolter TD: Accuracy of patient prediction of allergy skin test results. Ann Allergy Asthma Immunol 2000, 85:382–384.
5. Williams PB, Ahlstedt S, Barnes JH, et al.: Are our impressions of allergy test performances correct? Ann Allergy Asthma Immunol 2003, 91:26–33.
6. Bousquet J, Lebel B, Dhlvert H, et al.: Nasal challenge with pollen grains, skin-prick tests and specific IgE in patients with grass pollen allergy. Clin Allergy Immunol 1987, 17:529–536.
7. Petersson G, Dreborg S, Ingestad RL: Clinical history, skin prick test and RAST in the diagnosis of birch and timothy pollinosis. Allergy 1986, 41:398–407.
8. Go AS. Refining probability: an introduction to the use of diagnostic tests. In: Friedland DJ, ed. Evidence-Based Medicine: A Framework for Clinical Practice. Stamford, CT: Appleton and Lange; 1998:11-34.
9. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA. 1994;271:703-7.
10. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. BMJ. 1994; 308: 1552.
11. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ. 1994; 309: 102.
12. Shi R, Conrad SA. Correlation and regression analysis. Ann Allergy Asthma Immunol. 2009; 103(4 Suppl 1):S35-41.
13. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem. 1993;39: 561-77.
14. Barnard GA, Studies in the History of Probability and Statistics: IX. Thomas Bayes' Essay Towards Solving a Problem in the Doctrine of Chances, Biometrika 45:293–295, 1958.

15. Katz MA. A probability graph describing the predictive value of a highly sensitive diagnostic test. N Engl J Med. 1974; 291: 1115-6.
16. Croskerry P. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Adv Health Sci Educ Theory Pract. 2009; (14 Suppl 1): 27-35.
17. Croskerry P. A universal model of diagnostic reasoning. Acad Med. 2009; 84:1022-8.
18. Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf. 2013; 22 (Suppl 2): ii21-ii27.
19. Croskerry P, Singhal G, Mamede S. Cognitive debiasing 1: origins of bias and theory of debiasing. BMJ Qual Saf. 2013; 22 (Suppl 2): ii58-ii64.
20. Ledley RS, Lusted LB. Probability, Logic and Medical Diagnosis. Science. 19599; 130: 892-930.
21. Lusted LB. Decision-making studies in patient management. N Engl J Med. 1971; 284: 416-24.
22. Eddy DM, Clanton CH. The art of diagnosis: solving the clinicopathological exercise. N Engl J Med. 1982; 306: 1263-8.
23. Bates DW, Gawande AA. Error in medicine: what have we learned? Ann Intern Med. 2000; 132: 763-7.
24. Charlin B, Gagnon R, Pelletier J, et al. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. Med Educ. 2006; 40: 848-54.
25. Epner PL, Gans JE, Graber ML. When diagnostic testing leads to harm: a new outcomes-based approach for laboratory medicine. BMJ Qual Saf. 2013; 22 (Suppl 2): ii6-ii10.
26. Croskerry P. Cognitive forcing strategies in clinical decision making. Ann Emerg Med. 2003; 41: 110-20.
27. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003; 78: 775-80.
28. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. Arch Intern Med. 2005; 165: 1493-9.
29. Federspil G, Vettor R. Clinical and laboratory logic. Clin Chim Acta. 1999; 280: 25-34.
30. Licata G. Probabilistic and fuzzy logic in clinical diagnosis. Intern Emerg Med. 2007; 2 :100-6.
31. Kuhn GJ. Diagnostic errors. Acad Emerg Med. 2002; 9: 740-50.

32. Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. BMJ. 2000; 320: 774-7.
33. Dovey SM, Meyers DS, Phillips RL Jr, Green LA, Fryer GE, Galliher JM, Kappus J, Grob P. A preliminary taxonomy of medical errors in family practice. Qual Saf Health Care. 2002; 11: 233-8.
34. Di Lorenzo G, Leto-Barone MS, La Piana S, La Porta G, Montalto G. Assessment of a qualitative serological assay to screen for allergic sensitization in elderly subjects. Allergy Asthma Proc. 2013; 34: e9-13.
35. Heit E, Rotello CM. Relations between inductive reasoning and deductive reasoning. J Exp Psychol Learn Mem Cogn. 2010 May;36(3):805-12.
36. Pastorello EA, Incorvaia C, Ortolani C, Bonini S, Canonica GW, Romagnani S, Tursi A, Zanussi C. Studies on the relationship between the level of specific IgE antibodies and the clinical expression of allergy: I. Definition of levels distinguishing patients with symptomatic from patients with asymptomatic allergy to common aeroallergens. J Allergy Clin Immunol. 1995; 96: 580-7.
37. Reid MC, Lane DA, Feinstein AR. Academic calculations versus clinical judgments: practicing physicians' use of quantitative measures of test accuracy. Am J Med. 1998; 104: 374-80.
38. Doust J. Diagnosis in General Practice. Using probabilistic reasoning. BMJ. 2009;339: b3823.
39. Heneghan C, Glasziou P, Thompson M, Rose P, Balla J, Lasserson D, et al. Diagnostic strategies used in primary care. BMJ. 2009; 338: b946.
40. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science. 1981; 211: 453-8.
41. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science. 1974; 185: 1124-31.
42. Redelmeier DA, Koehler DJ, Liberman V, Tversky A. Probability judgement in medicine: discounting unspecified possibilities. Med Decis Making. 1995; 15: 227-30.
43. Lee S, Hess EP, Nestler DM, Bellamkonda Athmaram VR, Bellolio MF, Decker WW, Li JT, Hagan JB, Manivannan V, Vukov SC, Campbell RL. Antihypertensive medication use is associated with increased organ system involvement and hospitalization in emergency department patients with anaphylaxis. J Allergy Clin Immunol. 2013;131: 1103-8.

44. Redelmeier DA, Shafir E, Aujla PS. The beguiling pursuit of more information Med Decis Making. 2001; 21:376-81.
45. Cosmancini G. La medicina non è una scienza. Breve storia delle sue scienze di base. Pag 1-121. Raffaello Cortina Editore.
46. Settipane RA, Schwindt C. Chapter 15: Allergic rhinitis. Am J Rhinol Allergy. 2013; 27 (Suppl 1): S52-5.
47. Settipane RA, Kaliner MA. Chapter 14: Nonallergic rhinitis. Am J Rhinol Allergy. 2013; 27 (Suppl 1): S48-51.
48. Di Lorenzo G, Pacor ML, Amodio E, Leto-Barone MS, La Piana S, D'Alcamo A, Ditta V, Martinelli N, Di Bona D. Differences and similarities between allergic and nonallergic rhinitis in a large sample of adult patients with rhinitis symptoms. Int Arch Allergy Immunol. 2011;155: 263-70.
49. Gendo K, Larson EB. Evidence-based diagnostic strategies for evaluating suspected allergic rhinitis. Ann Intern Med. 2004 Feb 17;140(4):278-89.
50. Wang DY, Chan A, Smith JD. Management of allergic rhinitis: a common part of practice in primary care clinics. Allergy. 2004 Mar;59(3):315-9.
51. Borghesan F, Bernardi D, Plebani M. In vivo and in vitro allergy diagnostics: it's time to re-appraise the costs. Clin Chem Lab Med. 2007;45: 391-5.
52. Pantin CF, Merrett TG. Allergy screening using a microcomputer. Br Med J (Clin Res Ed). 1982; 285: 483-7.
53. Chae YM, Jang TY, Park IY, Chung SK, Park M. The development of a decision support system for diagnosing nasal allergy. Yonsei Med J. 1992; 33: 72-80.
54. Lindberg S, Malm L. Comparison of allergic rhinitis and vasomotor rhinitis patients on the basis of a computer questionnaire. Allergy. 1993; 48: 602-7.
55. Chae YM, Ho SH, Chung SK, Hong CS, Park IY. Structural modeling of differential diagnosis, treatment, and results for allergic rhinitis. Yonsei Med J. 1995; 36:116-29.
56. Park KS, Chae YM, Park M. Developing a knowledge-based system to automate the diagnosis of allergic rhinitis, Biomedical Fuzzy and Human Sciences. 1996; 1: 9–18.
57. D'Amato G, Dal Bo S, Bonini S. Pollen-related allergy in Italy. Ann Allergy. 1992; 68: 433-7.
58. Heinzerling L, Frew AJ, Bindslev-Jensen C, Bonini S, Bousquet J, et al. Standard skin prick testing and sensitization to inhalant allergens across Europe--a survey from the GALEN network. Allergy. 2005; 60: 1287-300.

59. Armitage P, Berry G. Statistical Methods in Medical Research (3rd edition). Blackwell 1994.
60. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley 1989.
61. Zhou X, Obuchowski NA, McClish DK. Statistical Methods in Diagnostic Medicine. New York: Wiley 2002.
62. Fletcher RW and Fletcher SW. Clinical epidemiology: The essentials, fourth ed., Lippincott Williams & Wilkins, 2005.
63. Lacagnina V, Leto-Barone MS, La Piana S, La Porta G, Pingitore G, Di Lorenzo G. Comparison between statistical and fuzzy approaches for improving diagnostic decision making in patients with chronic nasal symptoms. Pubblished online http://www.sciencedirect.com/science/article/pii/S0165011413004405