Exhaled air temperature in asthmatic children: a mathematical evaluation

Faster rise of exhaled breath temperature was reported as a marker of airway inflammation (1, 2) in adult asthmatics by evaluating the rate of temperature increase (Δe°T) between the beginning of exhalation and 63.2% of the total increase (3). This parameter was found to be significantly higher in both intermittent and persistent asthmatics than in controls and to positively correlate with exhaled nitric oxide (eNO) in patients with asthma.

During another study performed in allergic asthmatic children, a significant relationship between the plateau of exhaled air temperature (PLET) and eNO was observed (4). Furthermore, house dust mite (HDM) avoidance in a high-altitude, low-humidity environment was associated, in HDM-sensitive asthmatic children, with both a reduction in sputum eosinophil cell counts and PLET, suggesting a possible relationship between exhaled air temperature and airway inflammation (5).

However, it was argued that PLET cannot be separated in the tracing of exhaled temperature from the peak of exhaled temperature and that PLET does not distinguish asthmatics from healthy controls (6). As no healthy children have been previously studied, the aim of this study was to verify the ability to distinguish asthmatics from normal controls by a dedicated detailed mathematical evaluation of the exhaled air curve under controlled conditions of temperature and humidity in a large number of children.

Methods

Patients

Ninety children (50 male, 40 female) with mild or moderate asthma (7) ranging in age between 9 and 16 yr (mean ± s.d. = 12.6 ± 2.25), with no exacerbations and no systemic steroid treatment in the previous 2 months were enrolled in the study.

All of them received the appropriate controller treatment according to the severity of the disease (8), i.e. low-dose inhaled corticosteroids, administered...
with an add-on therapy with long-acting $\beta_2$-agonists or a leucotriene receptor antagonist for the group with moderate persistent disease.

Thirty-three healthy controls (18 male, 15 female) ranging in age between 12 and 13 yr (mean $\pm$ s.d. = 12.45 $\pm$ 0.50) were also evaluated in the same period.

No subjects with cardiac or lung malformation, with current infectious febrile diseases or with metabolic diseases, were not included in the study. Neither patients nor controls were smokers.

Study design

Children were evaluated in the morning and exhaled breath temperature was measured after at least 1 h of rest in the laboratory where the temperature in the range 20–22°C and the humidity between 45% and 55% were always maintained. During this time, children were asked to avoid running or doing any vigorous play. Body temperature was always measured at the ear using an instant thermometer (ThermoScan pro LT, Type 6007, professional model, Braun; ThermoScan Inc., San Diego, CA, USA) and exhaled air temperature was evaluated during a slow expiratory manoeuvre from total lung capacity. Three measurements were performed allowing 10 min of rest between evaluations to avoid interference caused by repeated efforts. This method allowed a training effect, which may have contributed to better performance of children. Therefore, only measures obtained by the last manoeuvre were considered for statistical analysis and mathematical evaluation. Long-acting $\beta_2$-agonists were stopped 24 h and short-acting 12 h before evaluation to avoid any influence of vasoactive medications on exhaled air temperature. The diet of asthmatic and control children was similar and breakfast was avoided in the morning of the study.

Approval for the study was obtained from the Local Hospital Ethical Committee and written consent was obtained from the parents of all the children.

Exhaled breath temperature measurement equipment and evaluation

The temperature of the exhaled air was evaluated using a COCO-001-T thermocouple (Unsheathed Fine Gage T/C; Omega Engineering Inc., Stamford, CT, USA) placed immediately after the sensor. In these conditions, it was possible to reach a plateau in exhaled air with a flow of 90 $\pm$ 10 ml/s. The COCO-001-T thermocouple is one of the commercially available thermocouples with the lowest thermal inertia, resulting in a response time (i.e. the time required to register 63.2% of an instantaneous change in temperature) of 45 ms. The sensor was placed within a T plastic tube between the mouthpiece and the flow restrictor. Digital data were sampled with a frequency of 200 Hz and 12 bits (error $<0.005^\circ$C/least significant bit) and were transmitted to a personal computer for the evaluation and storage of the data.

By means of a dedicated software, which was developed for the real-time display of the curve, it was possible to characterize at least four main phases of the expiratory air temperature pattern. Part A and part D corresponded to the phases of approach and separation from the mouthpiece (Fig. 1). The rapid increase in temperature at the beginning of exhalation was described by the zone B of the tracking. The intermediate part of the curve (zone C) was the segment where the PLET was located.

Statistical analysis

The mean value of temperature at the plateau (mvPLET) and the rate of the exhaled air temperature increase ($\Delta e^\circ T$) were determined. mvPLET was calculated on the mean of different values in the part of the curve where the temperature was included within a range of 0.5$^\circ$C at least for 2 s.

$\Delta e^\circ T$, expressed as $^\circ$C/s, was calculated between the beginning of exhalation and 63.2% of the total temperature increase (Fig. 2).

A statistical analysis of mvPLET and $\Delta e^\circ T$ was performed for the two groups of subjects by Student’s $t$-test. Data are reported as mean $\pm$ s.d. Significance was defined as a p-value of $<0.05$.

Mathematical evaluation

A pattern recognition analysis (9) was performed throughout the evaluation of several samples of exhaled air temperature for each subject. This allows a better definition of the phenomena than the simple evaluation of the two variables mvPLET and $\Delta e^\circ T$.

In particular, in our study, we considered for any subject 150 values of temperature obtained from the zone B (point of rapid rise in temperature) and zone C (plateau) of the curve (Fig. 2). Successively, to reduce the number of variables,
a separate analysis of 50 points from zone B and 100 points in zone C was performed in each single subject.

The principal component analysis (PCA) (10) was employed to classify the population under study by means of single variables and mvPLET and $\Delta e^o T$. PCA is a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables, which are ordered by reducing variability, called principal components. The same data were then evaluated using the technique of artificial neural networks (ANNs) (11, 12).

Principal components analysis

The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The uncorrelated variables are linear combinations of the original variables, and the last of these variables can be removed with minimum loss of real data. The aim of PCA was to discover or to reduce the dimensionality of the data set and to identify new meaningful variables.

Without loss of generality, let us assume that input data matrix $X$ has zero empirical mean. We want to find an orthonormal projection matrix $P$ such that:

$$Y = P^TX$$

where $Y$ matrix is the matrix of the principal components and with the constraint that the covariance matrix of the principal components, $\text{cov}(Y)$, is a diagonal matrix and $P^{-1} = P^T$. By substitution and matrix algebra, it can be demonstrated:

$$\text{cov}(X) = P \text{cov}(Y) P^T$$

The most used mathematical method to calculate the orthonormal projection matrix $P$ is the
singular value decomposition (SVD). Defining a general \( N \times M \) matrix as \( A \), the SVD theorem states:

\[
A = U \Sigma V^T
\]

where \( U \) is an \( N \times M \) unitary matrix whose columns are the left singular vectors; \( \Sigma \) is an \( M \times M \) diagonal matrix of singular values; and \( V^T \) is a \( M \times M \) unitary matrix whose rows are the right singular vectors. \( U \) elements (vectors) form an orthonormal set, \( V^T \) elements (vectors) form an orthonormal set, and \( \Sigma \) is diagonal.

Considering as orthonormal set, i.e. \( P \) matrix elements, the eigenvectors of \( \text{cov}(X) \), it follows that the matrix \( \text{cov}(Y) \) becomes diagonal, which elements are the eigenvalues of \( \text{cov}(X) \):

\[
\text{cov}(Y) = \begin{bmatrix}
\lambda_1 & 0 & \ldots & 0 \\
0 & \lambda_2 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & \lambda_m
\end{bmatrix}
\]

Thus, the SVD, finding the eigenvalues and eigenvectors of the covariance matrix of data, i.e. the diagonalization of the covariance matrix, represents an expansion of the original data in a coordinate system where the covariance matrix is diagonal. The orthonormal projection matrix \( P \) is also called score matrix.

Artificial self-organizing neural networks

The concept of ANNs is to imitate the structure and workings of the human brain by means of mathematical models (13). ANNs possess an adaptable knowledge that is distributed over many neurons and synaptic connections. The structure of the single neuron model, the network topology and the adaptation strategy (learning rule) define the ANN architecture. The neurons (processing units) are single elements and consist principally of a connection function, an input function, an activation (transfer) function, and an output function. A neuron receives signals via several input connections. These are weighted at the input to a neuron by the connection function. The weights define the coupling strength (synapses) of the respective connections and are established via a learning process, in the course of which they are modified according to given patterns and a learning rule. In the case of supervised learning, in addition to the input patterns, the desired corresponding output patterns also presented to the network in the training phase. In the case of unsupervised learning, the network is required to find classification criteria for the input patterns independently. Stochastic learning methods employ random processes and probability distributions to minimize a suitably defined energy function of the network. A large number of neural models now exist, and each of these models is available in various forms. The Integrand-and-Fire (IF) neuron model is often used to implement ANNs suitable for classification and forecast tasks (14). The multi-layer perceptron (supervised learning), the self-organizing map (unsupervised learning) and the Kohonen self-organizing map (KSOM) (15) are examples of ANN architectures based on the IF neuron model. For the purpose of this study, KSOMs were applied.

A KSOM maps the input data into a two-dimensional net of artificial neurons to solve classification tasks and to find structures in data. Data set is partitioned in a training data set and a test data set. Firstly, in the unsupervised training process, the synaptic weight vectors of the artificial neurons of the KSOM are adapted by means of the training data set examples in such a way that the KSOM supplies as good a representation as possible of the training data set. The synaptic weight vector of an artificial neuron of a KSOM corresponds to the feature vector of an object in the feature space under study. During the training process of a KSOM, a winner-takes-it-all training algorithm is performed. For each \( m \)-dimensional input vector belonging to the training data set \( \vec{f}_k = f_{k,1}, \ldots, f_{k,m} \), the generic artificial neuron \( i \) that has the minimum distance \( d_i = \| \vec{f}_k - \vec{w}_i \| \) from the input vector and the synaptic weight vector \( \vec{w}_i \), is the winning unit \( z \). During each time step \( t \), an input vector \( \vec{f}_k \) is randomly selected from the training data set. Epoch time \( T \) is defined as the \( t \) time steps in which all the input vectors are selected. The training process goes on for a finite number of epochs. The weight \( w_{ij} \) of the synapse connecting the \( j \)th element of the randomly selected input vector, \( f_{k,j} \), with the \( i \)th artificial neuron at the time step \( t \) and epoch \( T \) is modified as follows:

\[
w_{ij}(t) = w_{ij}(t - 1) + \alpha(T) r_i(T) \left[ f_{k,j}(t) - w_{ij}(t - 1) \right]
\]

where:

- \( \alpha(T) = f_\alpha(T - 1) \), learning rate factor with a learning rate factor \( f_\alpha \). The learning rate factor is in the range \( ]0,1[ \).
\( r_i(T) = \exp(-D_i^2/\sigma^2) \), feedback function of \( i \)th neuron. \( D_i \) determines the Euclidean distance of neuron \( i \)th to the winning neuron \( z \) on the KSOM.

\( \sigma(T) = f_\sigma(T - 1) \), learning radius with learning radius factor \( f_\sigma \). The learning radius factor is in the range \( [0,1] \).

The learning rate and the learning radius thus fall off exponentially during training so as to bring the artificial neurons to a state of equilibrium.

After the training process, a second supervised labelling step is performed. Cluster labels are assigned to the individual artificial neurons. This is performed via the interpretation of the content of the synaptic weight vectors (feature vectors) of the artificial neurons. Here, the same label can be assigned to several artificial neurons. Thus, a cluster is represented by several artificial neurons in the KSOM. After validation of the KSOM by examples of test data set, performance of the classification task is commonly evaluated using the confusion matrix. The generic element \( c_{ij} \) of the confusion matrix indicates how many times in percentage, a pattern belonging to the class \( i \) was classified as belonging to the class \( j \).

Three different KSOM architectures, \( 4 \times 4 \), \( 5 \times 5 \) and \( 10 \times 10 \) artificial neurons, were trained and tested. For each architecture, we fixed \( \alpha(0) = 0.999 \), \( f_\alpha = 0.99 \), \( \sigma(0) = 5 \), \( f_\sigma = 0.995 \) parameters and a training of 10,000 epochs.

To check the classification capability of the KSOM, data sets belonging to B and C zones were analysed. To check the generalization capability of the neural network, a \( K \)-fold cross-validation was carried out. Cross-validation is one of several approaches to estimating the performance of a model on future as-yet-unseen data. In \( K \)-fold cross-validation, the original data set is partitioned into \( K \) subsets. For each cross-validation step, a single subset is retained as the test set, and the remaining \( K - 1 \) subsubsets are used as the training set. The cross-validation process is then repeated \( K \) times, with each of the \( K \) subsets used exactly once as the test set. The \( K \) results from the folds then can be averaged (or otherwise combined) to produce a single estimation. In this work, a threefold cross-validation was applied; each fold consisted of randomly selected 11 people with asthma and 30 control subjects. Three different KSOM architectures, \( 5 \times 5 \), \( 10 \times 10 \) and \( 15 \times 15 \) artificial neurons, were trained and tested.

### Results

Body temperature of the children ranged between 36.6 and 37.0°C without any statistically significant differences between asthmatic children and controls.

Means and standard deviations of \( \text{mvPLET} \) and \( \Delta e^\circ T \) are reported in Table 1. As can be seen, \( \text{mvPLET} \) was significantly higher (\( p < 0.0001 \)) in asthmatic children than in controls, whereas this was not the case for \( \Delta e^\circ T \) (Fig. 3).

The results of the PCA are expressed in Fig. 4. Fig. 4a represents the ‘circle of correlation’ which indicates the contribution of the single variable to the final information. According to the arrow direction shown in Fig. 4a, the variables are reported with increasing values.

<table>
<thead>
<tr>
<th></th>
<th>Asthmatic children</th>
<th>Control subjects</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{mvPLET (°C)} )</td>
<td>31.15 ± 1.19</td>
<td>30.27 ± 1.25</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>( \Delta e^\circ T (\circ C/s) )</td>
<td>116.43 ± 87.05</td>
<td>120.90 ± 95.90</td>
<td>0.83</td>
</tr>
</tbody>
</table>

\( \text{mvPLET} \), mean value of temperature at the plateau. \( \Delta e^\circ T \), rate of temperature increase.

**Table 1.** Average and the standard deviation values for asthmatic children and control subjects

**Fig. 3.** (a) Mean value of temperature at the plateau (\( \text{mvPLET} \)) in asthmatic children vs. control subjects. (b) Rate of temperature increase (\( \Delta e^\circ T \)) in asthmatic children vs. control subjects.
starting from the first value selected in the zone of rapid rising of temperature up to the last value selected in the plateau area. In this system, the values which are closer to 1.0 give the highest contribution to the information. As can be seen, the major share to the information is given by the data taken from the plateau area (Fig. 4a) and the two populations (asthmatic children and healthy controls) are clearly separated into two clusters (Fig. 4b) even if there is a partial overlap.

To reduce the number of variables, so that it is possible to understand better which part of the curve gives the larger contribution to the information, 50 points in zone B (rapid rising of temperature) (Fig. 5a) and 100 points in zone C (plateau) (Fig. 5b) were evaluated by PCA. Again, the major contribution to the information comes from the values selected in the plateau area.

The results of the ANN are reported in Fig. 6. A topological analysis of the KSOM’s status after training showed the presence of two non-overlapped regions for each pattern of the training data set. To quantify results obtained in the threefold cross-validation, a labelling process allowed the two regions to be classified into class, i.e., people with asthma and control subjects. Table 2 summarizes the percentages of the mean confusion matrix that resulted from the averaged threefold cross-validation results of the three KSOM configurations. As mentioned above, the generic element $c_{ij}$ of the confusion matrix indicates how many times in percentage a pattern belonging to the class $i$ was classified as belonging to the class $j$.

**Discussion**

Data from this study show that mvPLET but not $Δe^oT$ is able to distinguish between asthmatics...
and healthy children. To reduce possible confounding factors, all the measurements were performed in the morning, after at least 1 h of rest in an environment with controlled air temperature and humidity. In particular, the timing of temperature assaying, i.e. in the morning for all the children, was decided when planning the experimental design of this study to rule out the possible interference of physical activity and that of circadian body temperature variations. The potential effect of such variations on exhaled breath temperature needs to be assessed with a dedicated study.

A possible explanation for the discrepancy between our results and those obtained by another group (3) may come in part from the faster time of response of the thermocouple we employed which results in a shorter phase of temperature rise at the beginning of exhalation and, as a consequence, in a greater \( \Delta e^0 T \) value. This technical difference, with a more advanced instrument in this study, could, in part, account for a different shape of the rising part of the curve, where \( \Delta e^0 T \) is calculated, therefore reducing the possibility of making a direct comparison of the data generated in these two studies. Nevertheless, the most innovative issue in this study is represented by the detailed evaluation of 150 points of the exhaled air curve by means of PCA. With this analysis, a clear separation of the population under study into two clusters, i.e. asthmatic and healthy children, came out only by the plateau phase of the exhaled temperature curve (Fig. 4b). Furthermore, it seems reasonable to speculate that the mathematical expression of the phenomenon by using the two values \( \Delta e^0 T \) and mvPLET is a too coarse simplification associated with a meaningful loss of information. The single analysis of \( \Delta e^0 T \) would suggest the expression of a mathematical model with a single exponential. However, our data suggest that the exhaled breath temperature curve is better represented by a complex function with the involvement of multiple exponentials. In this regard, the PCA, performed separately in 50 points of the faster rising part of the curve and in 100 points of the plateau, clearly indicates that the major information comes from the plateau (Fig. 5b). The evaluation of 100 points of the plateau resulted in a further better separation of asthmatics from controls than the simple use of the mvPLET.

A possible explanation for the discrepancy between our results and those obtained by another group (3) may come in part from the faster time of response of the thermocouple we employed which results in a shorter phase of temperature rise at the beginning of exhalation and, as a consequence, in a greater \( \Delta e^0 T \) value. This technical difference, with a more advanced instrument in this study, could, in part, account for a different shape of the rising part of the curve, where \( \Delta e^0 T \) is calculated, therefore reducing the possibility of making a direct comparison of the data generated in these two studies. Nevertheless, the most innovative issue in this study is represented by the detailed evaluation of 150 points of the exhaled air curve by means of PCA. With this analysis, a clear separation of the population under study into two clusters, i.e. asthmatic and healthy children, came out only by the plateau phase of the exhaled temperature curve (Fig. 4b). Furthermore, it seems reasonable to speculate that the mathematical expression of the phenomenon by using the two values \( \Delta e^0 T \) and mvPLET is a too coarse simplification associated with a meaningful loss of information. The single analysis of \( \Delta e^0 T \) would suggest the expression of a mathematical model with a single exponential. However, our data suggest that the exhaled breath temperature curve is better represented by a complex function with the involvement of multiple exponentials. In this regard, the PCA, performed separately in 50 points of the faster rising part of the curve and in 100 points of the plateau, clearly indicates that the major information comes from the plateau (Fig. 5b). The evaluation of 100 points of the plateau resulted in a further better separation of asthmatics from controls than the simple use of the mvPLET.

As a further step in the analysis, we used the self-learning model based on ANNs, a complex and flexible non-linear system with properties not found in other modelling systems. These properties include robust performance in dealing with noisy or incomplete input patterns, high fault tolerance and the ability to generalize from the input data (16). Models based on ANN are applicable in the diagnosis of different diseases as well as in waveform analysis of EKGs, EEGs and radiographic images (17). The mathematical evaluation of our data, with the use of ANNs, reinforces the possibility of distinguishing between asthmatics and healthy subjects by the evaluating exhaled air temperature in the plateau phase of the curve.
These findings are consistent with the hypothesis that the end of expiration is more directly reflecting the degree of inflammation occurring in small airways where a ventilation limitation has to be expected even during period of disease control (18).

Thus, the better performance obtained with the plateau phase of the curve may be assimilated to the improved performance of small airways' lung function testing in detecting airways' obstruction during asymptomatic periods in asthmatic children (19, 20) and reflects the importance of peripheral lung mechanics in asthma (21, 22).

In conclusion, the mathematical evaluation of the curve tracing exhaled air temperature confirms and further extends the possibility of using breath temperature in distinguishing asthmatics from healthy subjects.

References