Unsupervised study of plethysmography signals through DTW clustering

Thibaut Germain[1,2], Charles Truong[1,2], Laurent Oudre[1,2] and Eric Krejci[1,2]

***Abstract*— The study of plethysmography time series is crucial to better understand the breathing behavior of mice, in particular the influence of neurotoxins on the respiratory system. Current approaches rely on a few respiratory descriptors computed on individual breathing cycles that fail to account for the variety of breathing habits and their evolution with time. In this paper we introduce a new procedure for the automatic analysis of plethysmography signals. Our method relies on a new and robust segmentation of respiratory cycles and a DTW-based clustering algorithm to extract the most typical respiratory cycles (called reference sequences). We can then create a symbolic representation of any new recording by matching respiratory cycles to their closest reference sequence. This new representation is a visual and quantitative tool to assess the breathing behavior of mice and its evolution with time. Our method is applied to plethysmography signals collected on mice with two different genotypes and exposed to a neurotoxin.**

Clinical relevance **This article proposes a novel approach to study plethysmography data. Our algorithm is able to accurately extract clinically meaningful respiratory cycles and the associated ventilation patterns descriptors such as tidal volume and inhalation/exhalation duration. In addition, thanks to the associated symbolic representation of signals, the temporal evolution of respiration is easily quantified. This opens a new research path to study the often slowly evolving and subtle influence of neurotoxins on the respiratory system.**

I. INTRODUCTION

Measuring respiratory function of conscious and spontaneously breathing animals is essential in many clinical studies on drug effects on the respiratory system [1]. Double chamber plethysmography (DCP) [2] is composed of two hermetically isolated compartments where the animal's head is in one chamber, its thorax and its rear in the other one. It independently tracks the nasal airflow and thoracic movements of a constrained animal. DCP is a relevant approach to evaluate ventilation mechanics of the respiratory system, and it provides insights on both ventilation function and lung function [3].

A respiratory cycle is defined as an inhalation followed by an exhalation [4]. The beginning of each phase is physiologically identifiable:

- t_{in} corresponds to the beginning of the inhalation phase, when the mouse starts to breathe air in and the lungs volume is minimal,
- t_{out} corresponds to the beginning of the exhalation phase, when the mouse starts to breathe air out and the lungs volume is maximal.

[1]Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, F-91190, Gif-sur-Yvette, France

[2]Université de Paris, CNRS, Centre Borelli, F-75005 Paris, France

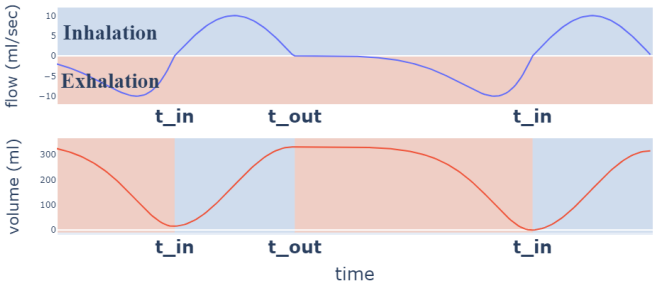

Fig. 1. Nasal airflow (top) and lung volume (bottom). During inhalation, airflow is positive (blue) and during exhalation, airflow is negative (pink). An inhalation starts at t_{in}, when the nasal volume is locally minimal. An exhalation starts at t_{out} when the nasal volume is locally maximal. A respiratory cycle starts at t_{in} and ends at the following t_{in}.

These phases and associated events are illustrated on Figure 1.

From plethysmography, respiratory function is commonly evaluated using several ventilation features (see [5], [6], [3] and Section III-B). While quantifying ventilation changes and their occurrences, these parameters provide few insights into physiological changes as they are often averaged on long periods of time. Recently, a new method based on machine learning has intended to leverage this issue [7] from whole-body plethysmography (WBP) signals, by using feature-based clustering of respiratory waveforms.

In this study, we introduce a new method for respiratory cycle analysis in the case of DCP data, which learns interpretable representations of typical respiratory cycles. These representations mirror distinctive breathing behaviors, and for a given recording, their succession in time provides insights into physiological changes. Our method is based on tools from machine learning for time-series: the K-Means clustering algorithm combined with the measure of fit Dynamic Time Warping (DTW). DTW compares time-series of different lengths and is invariant to a pattern's contraction or dilatation. This property is of particular interest to manage inter-individual variability. In addition, our algorithm relies on a new and robust algorithm to identify the starts and ends of inhalation and exhalation phases from a nasal airflow signal. As a significant difference from [7], our method evaluates the similarity between respiratory cycle patterns directly from shape variation of inhalations and exhalations rather than learned features, thus providing a more robust study of the respiratory physiological activity.

II. METHOD

Our method is composed of three main steps:

Step 1: Detection of the respiratory cycles and extraction of the inhalation/exhalation sequences. The first step of the process consists in extracting the respiratory

cycles from the input data. Each cycle is composed of two phases: inhalation and exhalation. A segmentation algorithm isolates the two periods. In a nutshell, given a raw signal $\mathbf{s}$, the first step of our method outputs a set of inhalation sequences $\{\mathbf{s}_{in}^{(1)}, ..., \mathbf{s}_{in}^{(N_s)}\}$ and a set of exhalation sequences $\{\mathbf{s}_{out}^{(1)}, ..., \mathbf{s}_{out}^{(N_s)}\}$, where N_s is the total number of cycles observed in the original signal $\mathbf{s}$.

Step 2: Computation of the reference sequences through an unsupervised clustering procedure. The second step consists in computing a small number of reference sequences from the sets of inhalation/exhalation sequences. The main idea behind this is to group sequences with common properties to highlight typical inhalation/exhalation behaviors. To that aim, a clustering algorithm (K-means) is used, combined with a measure of fit between signals that is able to compare sequences of different durations, namely the Time-Normalized DTW (TN-DTW), a robust variant of the DTW. The output of this step is a set of inhalation reference sequences $\{\mathbf{r_{in}}^{(1)}, \mathbf{r_{in}}^{(2)}, \dots\}$ and a set of exhalation reference sequences $\{\mathbf{r_{out}}^{(1)}, \mathbf{r_{out}}^{(2)}, \dots\}$.

Step 3: Characterization and symbolization of new recordings based on the extracted reference sequences. Consider a new recording $\mathbf{s}'$. The objective is to automatically characterize this recording using the reference sequences extracted in Step 2. First, the signal $\mathbf{s}'$ is segmented through the procedure already described in Step 1. Then, each of the $N_{s'}$ inhalation/exhalation sequences detected in $\mathbf{s}'$ is assigned a symbol that represents the reference sequence that is closest considering the measure TN-DTW.

A. Detection of the respiratory cycles and extraction of the inhalation/exhalation sequences

As mentioned previously, nasal airflow suffers from noise, making current inhalation and exhalation phases detection methods unreliable. Inaccurate detection then leads to biased descriptors and eventually to false experiments conclusions. We propose an algorithm that looks for local minima and maxima of the nasal volume (instead of airflow). Let $\mathbf{s}$ denote a raw nasal airflow.

First, the lung volume $\mathbf{v}$ is computed from the nasal airflow $\mathbf{s}$. This can be done by numerical integration:

$$v_t := \left(\sum_{u=1}^{t} s_u\right) - (\hat{a}t + \hat{b}) \tag{1}$$

where $\hat{a}, \hat{b} \in \mathbb{R}$ are such that $\sum_t v_t = 0$ and $\sum_t t v_t = 0$. The affine function $t \to \hat{a}t + \hat{b}$ removes the linear trend appearing during the integration process.

Next, the inhalation start times t_{in} and the exhalation start times t_{out} are identified using a peak-searching procedure that detects local minima, respectively maxima, of the nasal volume signal $\mathbf{v}$. To ensure an alternation between inhalation and exhalation, the algorithm first searches for all local minima (corresponding to the starts of the inhalations) and then searches for the maxima between two consecutive local minima. Once all inhalation/exhalation start times t_{in} and t_{out} are extracted, the original nasal airflow signal $\mathbf{s}$ is split into a set of inhalation sequences $\{\mathbf{s}_{in}^{(1)}, ..., \mathbf{s}_{in}^{(N_s)}\}$ and a set of exhalation sequences $\{\mathbf{s}_{out}^{(1)}, ..., \mathbf{s}_{out}^{(N_s)}\}$, where N_s is the total number of cycles observed in the original signal $\mathbf{s}$. The local minima search uses SciPy's (scipy.org) implementation; in our setting, only local minima t_{in} with a prominence above 0.03 mL in a window of 2 seconds are kept.

B. Computation of the reference sequences

Provided a set of inhalation/exhalation sequences, K reference sequences are computed. Each represents a typical respiratory behavior. In the following, $\mathcal{X} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}\}$ denotes a set of sequences (either inhalation or exhalation) of potentially different durations.

1) Clustering algorithm: The K reference sequences from the set $\mathcal{X}$ are computed with the well-known unsupervised clustering procedure called K-Means. This algorithm creates K non-overlapping groups $\{\mathcal{C}^{(1)}, \dots, \mathcal{C}^{(K)}\}$ of sequences with common properties. Roughly, K-Means is a two-step iterative refinement technique that assigns each sequence to the closest current centroid and then updates each centroid with regard to the new assignments. A centroid is a reference sequence $\mathbf{r}^{(i)}$ which corresponds to the average sequence of the cluster $\mathcal{C}^{(i)}$. Two crucial ingredients of the K-means algorithm are the measure of fit that is used to assign each sequence to a cluster and the procedure used to compute the reference sequences of each cluster. Although most publications usually use the Euclidean distance, this is not possible in our context since the sequences to cluster do not have the same duration. Also, the measure of fit must be invariant to some sequence properties: dilatation/contraction, amplitude offset, amplitude shift, time fluctuation, noise and outliers. For the experiments (Section III), the number of iterations of the K-means is set to 10.

2) Preprocessing: In order to be invariant to amplitude offset and shift, and to improve the robustness to outliers, all sequences are first centered to zero mean and scaled to unit variance. Then the discrete first order derivative is computed:

$$\Delta x_t = \begin{cases} x_2 - x_1 & \text{if } t = 1, \\ \frac{(x_t - x_{t-1}) + ((x_{t+1} - x_{t-1})/2)}{2} & \text{for } t = 1, ...T-1, \\ x_T - x_{T-1} & \text{if } t = T. \end{cases} \tag{2}$$

3) Time Normalized Dynamic Time Warping (TN-DTW): At each iteration, the K-Means algorithm assigns each sequence to the closest centroid. The distance is computed using TN-DTW which is a variant of DTW [8]. Both measures are invariant to time fluctuation and can compare sequences of different durations. Unlike DTW, TN-DTW is not affected by the length of the sequences. The relationship between these measures is given by:

$$\text{TN-DTW}(\mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{m+n}} \text{DTW}(\mathbf{x}, \mathbf{y}) \tag{3}$$

where $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$. DTW is computed in $O(mn)$ in time and space using dynamic programming. Computation time is reduced by using a Sakoe-Chiba constraint that prunes the set of possible matchings between samples [8]. This constraint is set to 20ms in the remaining of the article.

Intuitively, TN-DTW considers as highly similar (almost zero distance) two sequences of a given phenomenon occurring at different speeds. This property is of particular interest for our problem since some mice may inhale or exhale faster than others. To do so, TN-DTW finds an optimal mapping between a query sequence and a referent sequence by locally stretching or contracting the time axis. TN-DTW outputs the time normalized cumulative sum of the mapping pairwise distances. In addition to invariance to amplitude, shift and offset, the combination of the pre-processing and TN-DTW ensures robustness against time fluctuation, noise and outliers.

4) Time-series averaging: Finding an average sequence is an important sub-routines of K-Means algorithm. Indeed, the quality of each cluster is highly dependent on the quality of its centroid [9]. At each iteration, all sequences in the data set $\mathcal{X}$ are assigned to their closest centroids $\{\mathbf{r}^{(1)}, \ldots, \mathbf{r}^{(K)}\}$. Then, each centroid is updated with the average of its assigned sequences, computed as follows.

For any set $\{\mathbf{x}'^{(1)}, \ldots, \mathbf{x}'^{(N')}\}$ of N' sequences, the average, with respect to the TN-DTW, is the solution of the following optimization problem:

$$\underset{\mathbf{y}\in\mathbb{R}^L}{\operatorname{argmin}} \frac{1}{N'} \sum_{i=1}^{N'} \text{TN-DTW}^2(\mathbf{y}, \mathbf{x}'^{(i)}) \tag{4}$$

where $L > 0$ is the average duration of the sequences $\mathbf{x}'^{(i)}$.

Accurately and efficiently solving Problem 4 is not trivial. Traditional averaging methods cannot deal with the non-linear mapping between sequences of potentially different durations. A recent work [10] uses the subdifferentiability property of the function to optimize to develop a stochastic subgradient descent algorithm (S-DBA). For a trade-off between accuracy and speed, we implemented a batch version of S-DBA called BS-DBA.

C. Characterization and symbolization of new recordings

For a new recording $\mathbf{s}'$, we first perform the segmentation process described in Section II-B in order to extract the inhalation/exhalation sequences. Then, we use a 1-NN (nearest neighbor) algorithm to assign each sequence to the reference sequence, which is the closest to it, in the sense of the TN-DTW measure. This procedure yields a symbolic representation of $\mathbf{s}'$, where each respiratory cycle is replaced by a symbol composed of a letter (which specifies the type of inspiration) and a number (which specifies the type of expiration).

III. RESULTS AND DISCUSSION

A. Data

Our method is applied on a subset of data from an experiment that studies the impact of some neurotoxins (acetylcholinesterase (AChE) inhibitor) on the respiratory system of mice [5]. Mice of different genotypes were exposed to different AChE inhibitors. For each mouse, the manipulation was as follows:

1) Phase 1: The mouse is placed in a DCP for approximately 15 minutes to serve as a baseline.
2) Phase 2: The mouse is removed from the DCP and exposed to AChE inhibitor.
3) Phase 3: The mouse is placed back into the DCP and its breathing was recorded for 35 minutes.

Adult male and female mice were maintained on a mixed B6D2 genetic background. Before starting any experiment, they were housed for 7 days, in an environment maintained at 23 ± 0.5 °C, 38-41% humidity, under a 12-hour light/dark cycle, with light provided between 7 am and 7 pm. All experiments were performed in accordance with the Council of European Committees Directive (86/609/EEC) and were approved by the Paris Descartes University Ethics Committee for Animal Experimentation (CEEA34.EK/AGC/LB.111.12).

In this study, a total of 12 mice, 6 wild type (WT) and 6 with Colq KO (COLQ) genotype [11] are selected. All were exposed to the same AChE inhibitor: physostigmine. The train set is composed of 5 WT and 5 COLQ. The remaining is used as a test set for symbolization. All recordings are sampled at 2000Hz and only parts corresponding to Phases 1 and 3 of the manipulation are considered. The training set is composed of 4000 respiratory cycles: 200 from Phase 1 and 200 from Phase 3 are randomly sampled for each mouse. Respiratory cycles are detected with our procedure, see Section II-A.

B. Description of the reference inhalation/exhalation sequences

For all experiments, the number of reference sequences is 5 for both inhalation and exhalation sequences. Figure 2-a and Figure 2-b display the learned reference sequences for the inhalation and exhalation phases, respectively. As well as a letter or a number, a unique color is attributed to each cluster/references sequence. The length of each reference sequence is set to the average length of sequences within the associated cluster. Additionally, each reference sequence is scaled by the average standard deviation of sequences within the associated cluster. Each inhalation/exhalation cluster is described with common respiratory pattern descriptors to complete the analysis. To ease comparison, we used polar plots for visualization (Figure 2-c / Figure 2-d) and the descriptor data set is normalized before averaging for each cluster. All following descriptors are widely used in the literature [3]:

- Inhalation/exhalation time (Ti/Te, s): inhalation/exhalation duration,
- Peak inhalation/exhalation flow (Pif/Pef ml/s) : maximum/minimum flow during the inhalation/exhalation,
- Nasal tidal/exhalation volume (NTV/NEV, ml): nasal volume of air inhaled/exhaled during inhalation/exhalation,
- Active inhalation/exhalation (AI/AE, s): duration of the period which starts when the flow reaches 5% of the Pif/Pef and ends at the beginning of the next exhalation/inhalation,

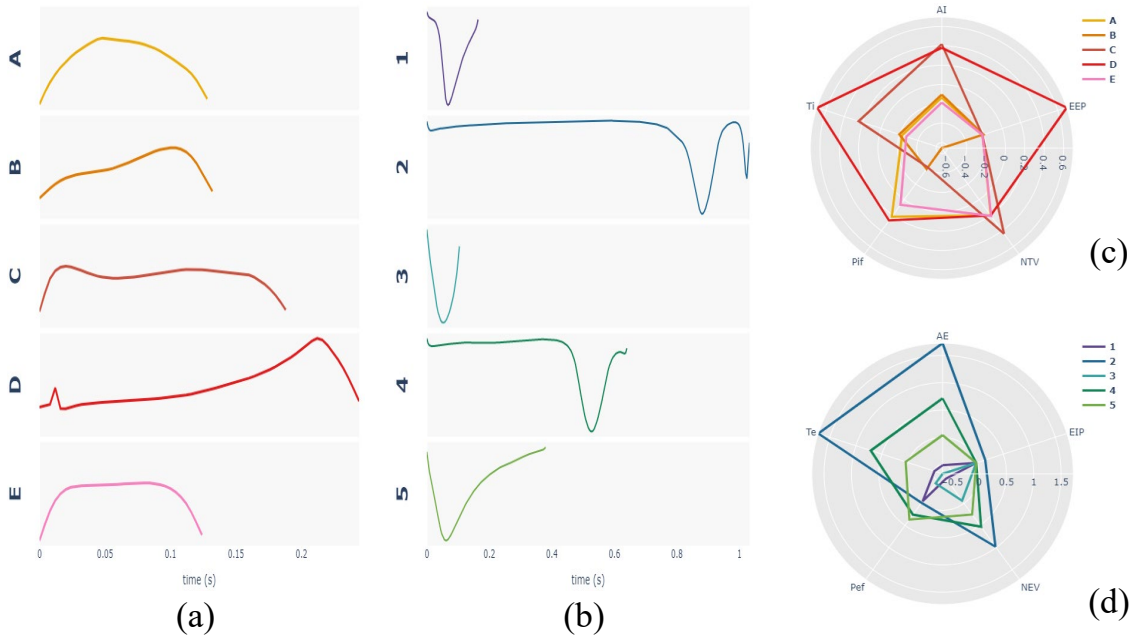

Fig. 2. Nasal airflow of the learned inhalation reference sequences (a) and exhalation reference sequences (b). (c) Inhalation cluster polar plot. (d) Exhalation cluster polar plot. Inhalation descriptors: active inhalation (AI), end exhalation pause (EEP), nasal tidal volume (NTV), peak inspiration flow (Pif), inhalation time (Ti). Exhalation descriptors: active exhalation (AE), end inhalation pause (EIP), nasal exhalation volume (NEV), peak exhalation flow (Pef), exhalation time (Ti).

- End inhalation/exhalation pause (EIP/EEP, s): duration of the period between the inhalation/exhalation start and active inhalation/exhalation start.

As shown on Figure 2-a, the learned inhalation reference sequences exhibit different behaviors. This fact is confirmed by the respiratory descriptors (Figure 2-c). For instance, Sequence `D` stands out when looking at both the graph and the descriptors. Indeed, it has a longer duration (0.25s) with a low airflow at the beginning and a maximum at the end, and has the largest Ti, EEP and Pif. (Recall that a low airflow means that the mouse has trouble breathing.) In contrast, Sequence `C` is shorter (less than 0.2s) but reaches its maximum faster and airflow remains high for most of the sequence. This is in accordance with the fact that it has in the largest NTV (inhaled volume) and AI (active inhalation) of all learned sequences. Sequence `B` is also striking (compared to sequences of same duration such as `A` and `E`) as it has a low airflow and reaches its maximum at the end, similarly to Sequence `D`. Unsurprisingly, the associated NTV is the lowest of the learned sequences. As for the exhalation sequences, two groups of reference sequences can be seen on Figure 2-b: those with a long pause (very low airflow) at the beginning (Sequences `2` and `3`) and those without (`1`, `3`, and `5`). The classical respiratory descriptors (Figure 2-d) cannot discriminate easily the two groups. They only notice that Sequences `2` and `3` are longer, but do not highlight the pause. To conclude, most of the learned reference sequences are associated with distinct behaviors according to the usual descriptors. Nevertheless, a few are not clearly distinguished by the descriptors, while our algorithm classify them as different, yielding a richer range of behaviors which the respiratory descriptors cannot describe.

C. Relevance of the symbolization process

Figure 3 provides a visualization of symbolic representation before and after drug injection for all mice present in the data set. For the sake of conciseness, only 10 minutes are displayed for Phase 1, and up to 25 minutes for Phase 3. Each bi-colored line is one mouse symbolization, where the upper line corresponds to the inhalations (warm colors) and the lower line to the exhalations (cold colors).

Such visualization allows for an overall understanding of the diversity of breathing behaviors and changes in breathing dynamics when facing stressors. In this experiment, Figure 3 shows, for instance, that each learned reference sequence is predominant for a specific genotype before or after drug injection.

Before injection, the breathing behavior of COLQ mice is characterized by the absence of pause: the respiratory cycle `C5` is predominant. Sequence `C` suggests a healthy inhalation: the volume of air inhaled is optimal (large NTV) while the effort is minimum (small Pif). After injection, `C` degrades to an alternation between periods of mostly `B` sequences and others of mostly `D` sequences. Both `B` and `D` suggest difficulties to inhale: `B` has a short duration (Ti) and a small inhalation volume (NTV), meaning that the mouse inhales superficially, and `D` has a long duration (Ti) and a pause before inhalation (EEP). Several physiological phenomena can cause such a behavior, e.g. a blocked diaphragm. After injection, exhalation Sequences `1` and `3` become frequent. Both have a short duration (Te), and a small exhalation volume (NEV) compared to `5`. Sequence `3` is prevalent for one mouse after injection (5^{th} COLQ mouse). The clustering algorithm has caught an individual behavior that is characteristic of this mouse.

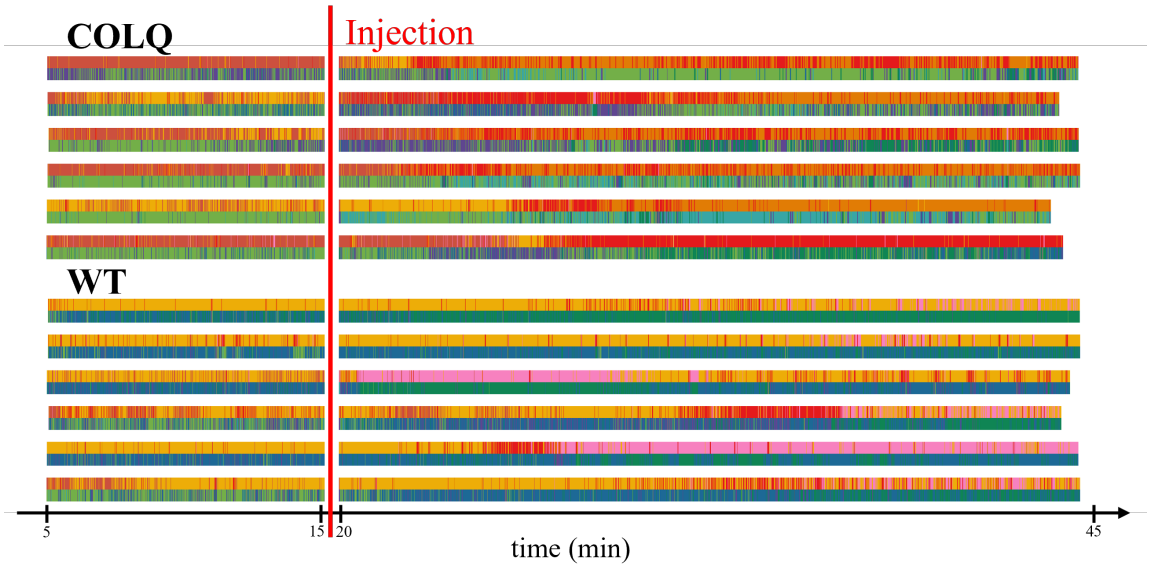

Fig. 3. Symbolic representation timeline. A bi-colored line is one mouse symbolic representation, the upper line corresponds to the inhalations (warm colors) and the lower line to the exhalations (cold colors). A color refers to the associated reference sequence in Figure 2. Respiratory cycles are stacked in timely order according to their length. 10 minutes before drug injection are symbolized and it goes up to 25 minutes. Last symbolization of each group corresponds to the test mouse.

WT mice breathing behavior is less sporadic compared to COLQ mice. Exhalation behavior is mainly characterized by Sequences `4` and `5`. Inhalation `A` and `C` are prevalent in all recordings. After injection, pauses start to appear (Sequence `D`). We can observe an alternation of inhalation with and without pause. As mentioned previously, Sequences `A` and `E` look similar based on their descriptors (Figure 2-c) but it is noticeable that while `A` occurs at any time, `E` is only present after drug injection. Compared to `A`, `E` reaches a plateau whose flow is lower than the maximum flow of `A` (Figure 2-a), suggesting that the drug injection induces the appearance of bronchoconstriction. It refers to inhaling difficulties due to the bronchus's constriction: the nasal airflow is limited. Pattern comparison suggests a significant physiological change that could have been difficult to catch with respiratory descriptors.

Both COLQ and WT develop a diversity of breathing behavior to the drug injection. Among others, we observed alternating periods composed of either long or short inhalation and exhalation. Studying breathing behavior changes with statistical analysis of respiratory descriptors may lead to biased results due to the fast-changing dynamic of behaviors. Indeed, taking the average inhalation or exhalation duration per minute would hide these alternations. The fast-changing dynamic could be analyzed more precisely by taking advantage of the symbolic representation.

IV. CONCLUSION

In plethysmography, the pattern comparison from the DTW-based clustering procedure complements the classical analysis: it highlights respiratory behaviors not detectable with the descriptors. The reference sequences display interpretable information and are related to ventilation parameters of well-established models in the literature. The symbolic representation paves the way for an accurate analysis of rapidly changing behaviors. The presented work could be extended to similar plethysmography data, in particular optical plethysmography, which is used to monitor human blood flow.

REFERENCES

[1] D. J. Murphy, "Assessment of respiratory function in safety pharmacology," *Fundamental & clinical pharmacology*, vol. 16, no. 3, pp. 183–196, 2002.

[2] H. Hoymann, "Lung function measurements in rodents in safety pharmacology studies," *Frontiers in Pharmacology*, vol. 3, p. 156, 2012.

[3] S. Mailhot-Larouche, L. Deschênes, K. Lortie, M. Gazzola, D. Marsolais, D. Brunet, A. Robichaud, and Y. Bossé, "Assessment of respiratory function in conscious mice by double-chamber plethysmography," *Journal of visualized experiments: JoVE*, no. 137, 2018.

[4] J. H. Bates and C. G. Irvin, "Measuring lung function in mice: the phenotyping uncertainty principle," *Journal of applied physiology*, vol. 94, no. 4, pp. 1297–1306, 2003.

[5] A. Nervo, A.-G. Calas, F. Nachon, and E. Krejci, "Respiratory failure triggered by cholinesterase inhibitors may involve activation of a reflex sensory pathway by acetylcholine spillover," *Toxicology*, vol. 424, p. 152232, 2019.

[6] R. Hill, R. Santhakumar, W. Dewey, E. Kelly, and G. Henderson, "Fentanyl depression of respiration: comparison with heroin and morphine," *British journal of pharmacology*, vol. 177, no. 2, pp. 254–265, 2020.

[7] M. D. Sunshine and D. D. Fuller, "Automated classification of whole body plethysmography waveforms to quantify breathing patterns," *Frontiers in Physiology*, p. 1347, 2021.

[8] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series.," in *KDD workshop*, vol. 10, pp. 359–370, Seattle, WA, USA:, 1994.

[9] S. Aghabozorgi, A. S. Shirkhorshidi, and T. Y. Wah, "Time-series clustering - a decade review," *Information Systems*, vol. 53, pp. 16–38, 5 2015.

[10] D. Schultz and B. Jain, "Nonsmooth analysis and subgradient methods for averaging in dynamic time warping spaces," *Pattern Recognition*, vol. 74, pp. 340–358, 2018.

[11] G. Feng, E. Krejci, J. Molgo, J. M. Cunningham, J. Massoulié, and J. R. Sanes, "Genetic analysis of collagen q: roles in acetylcholinesterase and butyrylcholinesterase assembly and in synaptic structure and function," *The Journal of cell biology*, vol. 144, no. 6, pp. 1349–1360, 1999.