6.1 Introduction

The prevalence of chronic diseases, in general, is rising due to an ageing population, as well as due to environmental and lifestyle changes. This is particularly true for respiratory diseases, such as chronic obstructive pulmonary disease (COPD), which is a progressive and irreversible disease that results in airflow limitation and significant extra pulmonary effects, which limit physical activities [123]. Physical activity (PA), defined as any bodily movement produced by skeletal muscles that requires energy expenditure (EE) [124], is known to be a relevant indicator of COPD patients' health status. Current COPD treatment guidelines strongly recommend pulmonary rehabilitation programs with the intent of increasing PA and maintaining the changes over time, including when patients are discharged to home [16]. Although learning to exercise and maintaining exercising is really difficult, this is possibly less of an issue for other effective behavioural changes like reducing the sedentary time. New methodologies assessing PA and a better insight into daily PAs of patients with COPD are needed to accurately characterize the disease and assist clinicians in designing targeted therapeutic strategies and personalized coaching programs that do not rely on a priori fixed objectives [51]. Since patients are not able to accurately self-report their physical activities [125], models describing PA should be learned in an unsupervised fashion without the need of user annotations. Moreover, they should be suitable to harmoniously integrate multiple low-level PA measures (such as intensity levels of the activities performed or step counts), and the associated physiological responses into meaningful descriptors [50]. With recent improvements in wearable sensor technologies, it becomes easier in daily life to acquire massive amounts of different sensors data. At the same time, however, it is difficult to combine and extrapolate meaningful information in absence of any supervision or annotation. In this work, measures of PA and physiological responses are presented as the letters composing the words that describe PAs. The co-occurrence of these words in different ways and proportions during the day creates groups of PA constructs describing the main topics that pervade the day of a patient, i.e., PA routines. Routines could also occur at different times and in different proportions across the day for different patients or subgroups of patients, thus characterizing their activity behaviour. In this work, we show that the low-level fusion of sensors data into "words" combined with topic modelling holds the promise of discovering new and clinically relevant insights. In particular, this work provides the following contributions. (1) We propose a methodology to create a vocabulary of meaningful words describing PA from a set of multimodal PA measures without the need of any supervision or parameter tuning. Using the generated vocabulary, we discover PA routines that pervade daily life of healthy control subjects and a subgroup of COPD patients pairwise matched for age, gender, and body mass index (BMI). (2) We infer the underlying PA routine structure of the dataset available (comprising more COPD patients) on day segments of 30 min and on the first 6 h at once describing, respectively, the temporal regularities of the multimodal PA measures, and the estimated time spent by the subjects under study in each of the routines. (3) Finally, we investigate for the first time whether discovered PA routines are related to disease severity, whether they can be deployed to cluster subjects of the matched dataset and, furthermore, to recognize from which group each assessed day comes from.