kg·m$^{-2}$; pre-obese, 25 to 29.99 kg·m$^{-2}$; or obese, $\geq$ 30 kg·m$^{-2}$) and by Global Initiative for Chronic Obstructive Lung Disease (GOLD) classifications (2007 [67] 1 to 4; and 2011 [9] A to D). GOLD 2011 classification (A to D) was based on the degree of airflow limitation (GOLD grades 1 to 4) and symptoms (mMRC dyspnoea grades 0 to 4).

## 4.2.2 Selection of waking hours recordings

Firstly, the data collected with the SenseWear Armband devices were exported in the form of Microsoft Excel spreadsheets with one minute resolution. The data contains information about the sleeping time and, in particular, each minute assessed is marked by the SenseWear software as "sleeping" or "not sleeping" [68]. Then, in order to reduce the variability of the data, only minutes coded as "not sleeping" were selected for analysis. If a minute was coded as "sleeping" but had an intensity value higher than 2.0 metabolic equivalents of task (METs), which is compatible with light intensity, this minute was considered as "not sleeping" since it is very unlikely that a subject present such a high intensity whilst sleeping.

## 4.2.3 Stratification of physical activity measures

The software SenseWear Professional versions 6.1 and 7.0 were used for data analysis, providing minute-by-minute energy expenditure (EE) and METs. These two measures were stratified according to different criteria (and the combination of them): intensity (e.g., very light, light or moderate-to-vigorous intensity), duration (e.g., bouts of activity), period of the day (e.g., before or after midday), frequency (e.g., number of bouts per day); and quantity (e.g., absolute numbers or percentage of total). A bout of activity was defined as a period of consecutive minutes with a minimum duration (e.g., 10 minutes) and in the same intensity. These stratifications were performed with Matlab R2012b (Mathworks Inc., USA) and led to 180 distinct variables referred to as features, which were used for clustering of patients. Table XVI in the appendix presents the 180 features used for cluster analysis.

## 4.2.4 Sample size calculation

The main analysis in our study was the identification of clusters based on physical activity data. To the best of our knowledge, currently there are no sample size calculation formulas for cluster analysis as performed in our study. Some authors have suggested that the minimal sample size to include in studies using cluster analysis should be no less than $2^k$ cases, preferably 5 x $2^k$, with K being the number of variables considered for analysis [69, 70]. In our study, only 3 variables (i.e., the 3 components from the principal component analysis) were used for clustering. Therefore, the minimal sample size in our study should be 40 subjects, which is actually far below the actual number of participants included (i.e., 1001 subjects). Furthermore, our sample size is much larger than that of most previous studies using cluster analysis in COPD,