

recognition algorithms ensured that the MET estimation was insensitive to noise and random motion artefacts [104]. For each minute, the device also recorded steps count, information about the sleeping status of a patient (0=awake, 1=sleeping), and posture (0=lying down, 1=not lying down). Night-time sleep was defined as sleep that occurs between 21:00 and 06:00 [105]. For this analysis we considered only MET, ST, GSR, SC, and Sleeping Status (SL) data within this time interval. The following night-time sleep measures were also derived from the sleep status information provided by the sensor: night sleeping time, number of nocturnal sleeping bouts and duration of nocturnal sleeping bouts. Sleeping bouts were defined as consecutive minutes marked by the sensor as sleeping. Participants who wore the device for at least four nights (two weekdays + Saturday + Sunday) were included [44]. In total 1059 patients with COPD and 66 healthy controls were included in the analysis. The median number of nights analysed per patient was six (four during weekdays, two during weekends), resulting in a total of 6446 valid nights assessed, of which 4335 (67.3%) were during weekdays and 2111 (32.7%) during weekends.

7.4.3 Topic models

Sensor data from 66 healthy subjects and 66 matched patients with COPD were used to create the vocabulary of words. METs data were divided into intensity categories (*IC*) using the thresholds proposed by the American College of Sports Medicine: very light intensity (*VL*), < 2.0 METs; light intensity (*L*), 2.0 to 2.9 METs; and moderate-to-vigorous intensity (*MV*), ≥ 3.0 METs [12]. Minutes marked by the sensor as sleeping and with METs < 2.0 formed a separated category named sleeping (*S*). Step counts were converted in a binary form depending on whether the participant performed steps in each assessed minute (0=no steps performed, 1=steps performed). Temperature and galvanic skin response data were first cleaned from missing values and outliers (i.e. temperature values outside the range [24-40 °C] and GSR values outside the range [0-8 μ Siemens]) and then, for each subject, were centred across the mean over multiple assessed nights. An example of data stream for a single patient night is shown in Figure 44.

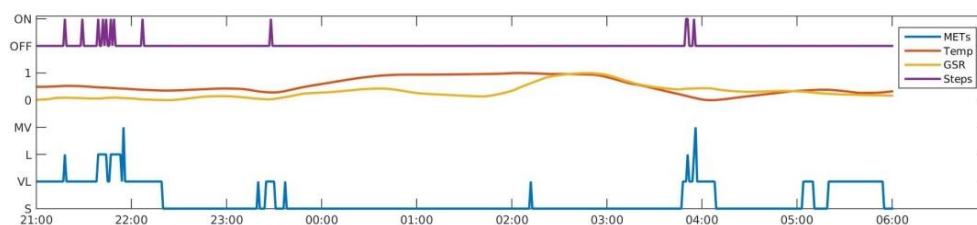


Figure 44 Night-time data from 21:00 pm to 06:00 am of a participant in the study. In blue Metabolic Equivalent of Task (MET), in yellow Galvanic Skin Response (GSR), in red Temperature (Temp), in purple Steps. METs data, defined as the energy cost of physical activities as a multiple of the resting metabolic rate, are divided in 4 intensity category (S, VL, L, and MV). GSR and temperature data are scaled in the interval 0-1 for visualization. Steps are discretized in a binary form (ON = at least 1 step performed, OFF = no steps performed).

In order to have sparse sleep modalities and symbols that best represent the original signal, it is desirable to have a discretization technique that produces symbols with equal probabilities [130] and that minimizes the distortion of the partitioned signal [155]. Therefore,