Supplementary materials for manuscript

A1. Data Collection

Everion Device

The Biovotion Everion device [1] has been selected for this study due to its extensive validation and use in several clinical trials and healthcare applications [2, 3]. The device is CE marked and has CE medical approval for HR and SpO2 parameters. Other technical details on the device can be found on the manufacturer's website [4].
On the practical side, the Everion device has battery life of approximately 24 hours; has multiple sensors in one device, which is vital for usability in real-world settings where multiple data sources should be acquired; provides high temporal resolution data (1 Hz); has an open API for incorporation in custom digital solutions for trials; has a comfort and discreet wear position.

Figure 1: Everion device (adapted from [1])

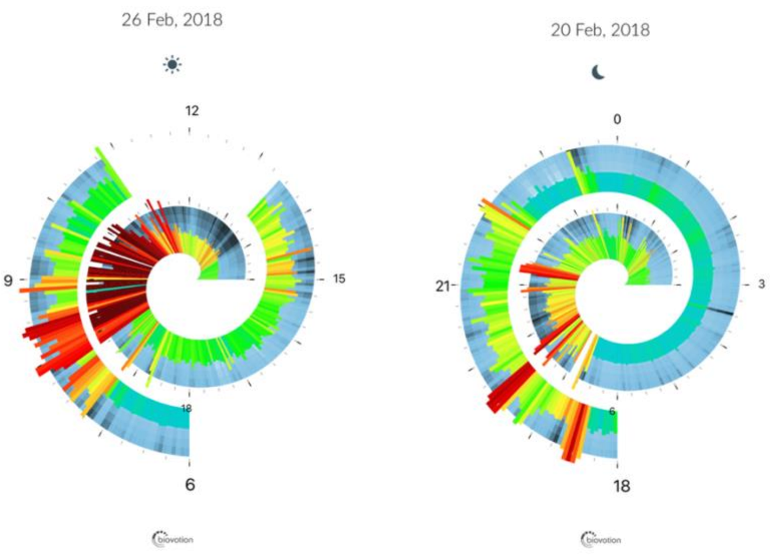

Figure 2: Example of screenshots from Biovotion X, the application used to interface with the Everion device.

SymTrack Application

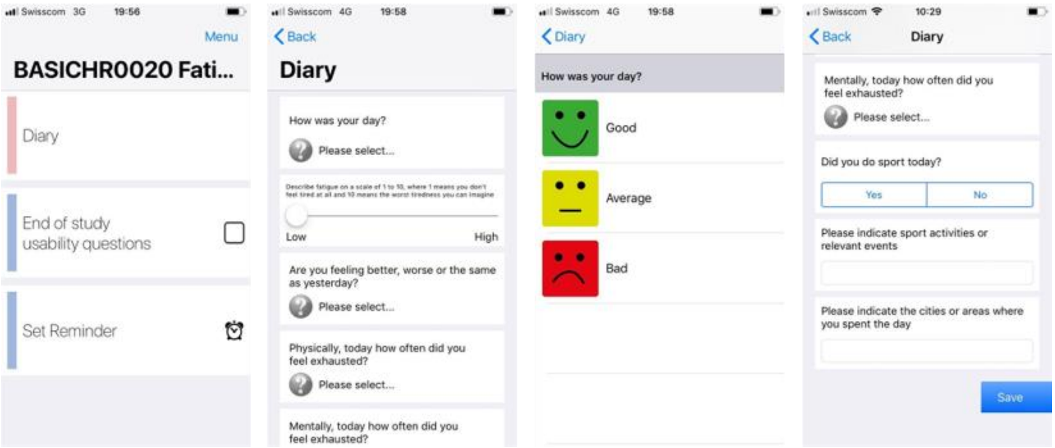

Figure 3: Screenshots of the SymTrack application, used to collect daily questionnaires on fatigue.

Data Statistical Summaries

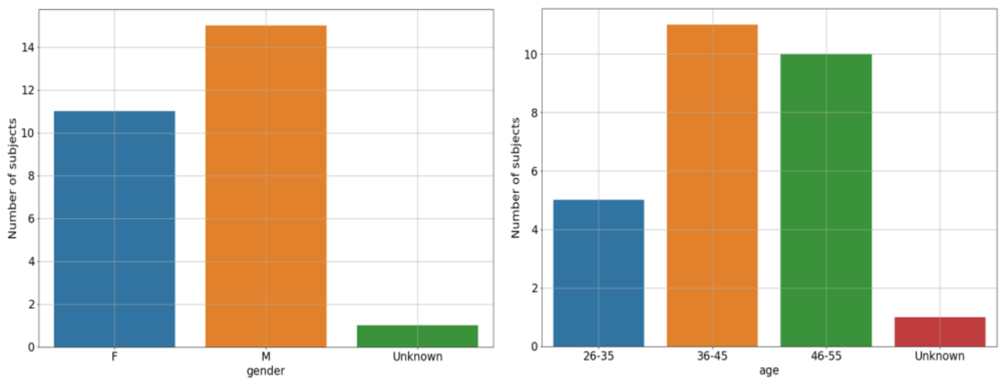

Figure 4: Demographics distribution (age and gender) of the 27 selected subjects.

Table 1: Subject-level demographics, duration of recordings and number of observations, and mean ± standard deviation (std) of reported fatigue scores. For the calculation of the Physical and Mental fatigue scores we have applied the following mapping: never = 0; sometimes = 1; regularly = 2; often 3; always = 4. PRO: patient reported outcomes (fatigue questionnaires); h: hours; M: male; F: female.

Subject No.	Age range	Gender	Original number of recording days	Number of recording days after pre-processing	Number of recording days with PROs	Jet lag (h)	Mean ± std Physical fatigue score	Mean ± std Mental fatigue score	Mean ± std VAS fatigue score
1	36-45	M	8	8	7	No	0.57 ± 0.53	0.43 ± 0.53	1.71 ± 0.76
2	36-45	M	1	0	0	No	1.43 ± 1.27	1.29 ± 1.25	4.86 ± 1.78
3	36-45	M	9	8	8	No	0.38 ± 0.74	1.00 ± 1.07	5.75 ± 1.83
4	36-45	F	10	9	7	No	0.57 ± 0.79	0.57 ± 0.79	3.29 ± 0.76
5	36-45	F	9	9	5	Yes (3.5)	1.00 ± 0.71	1.20 ± 1.30	5.00 ± 2.73
6	46-55	F	9	3	2	Yes (6)	1.00 ± 0.71	1.00 ± 0.71	7.40 ± 2.70
7	36-45	M	7	6	2	Yes (4.5)	0.50 ± 0.71	0.00 ± 0.00	1.50 ± 0.71
8	46-55	M	13	12	9	Yes (6)	1.33 ± 1.00	1.00 ± 1.00	3.67 ± 2.18
9	36-45	M	8	8	7	Yes (7)	0.43 ± 0.53	0.00 ± 0.00	7.86 ± 1.07
10	36-45	F	9	9	6	Yes (6)	0.50 ± 0.55	0.33 ± 0.52	2.00 ± 0.63
11	36-45	F	8	8	7	Yes (6)	1.14 ± 0.69	1.14 ± 0.90	5.29 ± 0.76
12	46-55	F	9	8	8	Yes (5)	1.25 ± 0.89	2.00 ± 0.93	6.13 ± 1.36
13	46-55	M	9	9	7	Yes (6)	0.29 ± 0.49	0.43 ± 0.53	1.71 ± 0.76
14	36-45	F	9	9	7	Yes (7)	1.29 ± 0.76	1.14 ± 0.90	3.14 ± 0.90
15	46-55	F	12	11	9	Yes (6)	0.78 ± 0.44	0.56 ± 1.01	1.67 ± 2.00
16	26-35	F	9	9	7	Yes (6)	0.71 ± 0.49	1.00 ± 0.58	2.57 ± 1.51
17	46-55	M	10	10	8	Yes (6)	0.38 ± 0.52	0.75 ± 0.89	1.63 ± 0.92
18	46-55	M	9	9	8	Yes (6)	0.25 ± 0.46	0.75 ± 0.71	2.38 ± 1.30
19	36-45	M	9	9	7	Yes (7)	1.43 ± 0.79	1.00 ± 1.00	4.00 ± 1.83
20	46-55	M	9	8	4	Yes (6)	1.75 ± 0.50	1.50 ± 0.58	6.50 ± 1.00
21	46-55	M	8	8	8	No	0.50 ± 0.53	0.00 ± 0.00	2.50 ± 0.76
22	unknown	unknown	13	12	7	unknow	0.71 ± 0.49	1.00 ± 0.58	2.57 ± 1.51
23	26-35	F	73	73	70	No	0.97 ± 1.18	0.86 ± 1.04	3.74 ± 2.28
24	26-35	M	217	130	97	No	1.16 ± 0.53	1.20 ± 0.78	4.52 ± 1.25
25	46-55	M	36	28	17	No	1.04 ± 0.98	0.70 ± 0.82	2.39 ± 0.58
26	26-35	M	209	66	47	No	1.51 ± 0.80	1.79 ± 1.06	5.04 ± 2.11
27	36-45	F	219	39	30	No	1.04 ± 1.00	0.48 ± 0.79	1.98 ± 1.09
28	26-35	M	22	12	4	No	0.20 ± 0.45	1.00 ± 0.71	4.60 ± 0.89
Total			973	530	405				

*Table 2: Subject-level mean ± standard deviation (std) of the 11 numerical sensor parameters and category ratios for the categorical parameter 'ActivityClass' over the first recording day (**baseline**). HR: heart rate; HRV: heart rate variability; RESP: respiratory rate; mbar: millibar; (m)s: (milli)seconds; kOhm: kilo Ohm; bpm: beats per minute; C: Celsius degree.*

Subject No.	Activity Counts	Barometer [mbar]	Blood Perfusion	Blood Pulse Wave	Energy Expenditure [calories/s]	Galvanic Skin Response [kOhm]	HR [bpm]	HRV [ms]
1	2.66 ± 4.32	990.11 ± 7.86	0.64 ± 0.27	2.31 ± 0.54	1564.82 ± 910.39	2.39 ± 1.05	72.44 ± 10.06	47.09 ± 11.35
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	1.92 ± 1.86	981.48 ± 9.65	0.47 ± 0.17	3.08 ± 0.68	1832.52 ± 1059.06	2.33 ± 0.87	83.61 ± 13.98	35.53 ± 6.45
4	1.68 ± 2.00	991.98 ± 2.55	1.21 ± 0.61	3.61 ± 1.19	1633.68 ± 1037.29	1.99 ± 0.98	87.83 ± 18.72	32.57 ± 13.74
5	2.02 ± 3.15	974.97 ± 2.25	0.24 ± 0.20	3.06 ± 0.86	1657.90 ± 908.48	2.57 ± 1.05	87.83 ± 14.11	42.37 ± 11.51
6	1.99 ± 1.83	990.24 ± 12.10	0.38 ± 0.15	2.26 ± 0.29	1592.76 ± 542.80	2.22 ± 0.25	70.55 ± 10.09	45.23 ± 9.52
7	1.79 ± 2.47	984.09 ± 4.65	0.40 ± 0.24	2.38 ± 0.57	1427.39 ± 772.65	1.97 ± 0.83	71.12 ± 13.73	45.27 ± 13.57
8	2.11 ± 3.74	970.27 ± 6.68	0.59 ± 0.41	2.37 ± 0.60	1436.93 ± 466.93	2.61 ± 0.68	68.27 ± 10.81	45.35 ± 9.94
9	2.03 ± 2.41	986.78 ± 5.45	0.56 ± 0.22	2.32 ± 0.83	1715.45 ± 870.34	2.50 ± 0.83	72.96 ± 17.76	33.42 ± 7.41
10	1.71 ± 1.93	980.57 ± 5.37	0.24 ± 0.18	2.32 ± 0.35	1465.63 ± 325.33	2.17 ± 0.81	72.59 ± 8.31	45.61 ± 9.28
11	1.63 ± 1.50	977.78 ± 5.61	0.36 ± 0.18	2.61 ± 0.45	1485.01 ± 308.71	1.22 ± 0.24	71.58 ± 7.00	46.75 ± 11.92
12	1.85 ± 2.19	978.86 ± 4.14	0.51 ± 0.23	2.78 ± 0.44	1395.86 ± 362.44	1.75 ± 0.79	74.15 ± 9.53	35.93 ± 10.63
13	2.09 ± 2.41	973.58 ± 18.15	0.70 ± 0.34	3.02 ± 0.82	1829.19 ± 1166.60	1.97 ± 0.91	82.85 ± 17.42	33.55 ± 11.55
14	3.88 ± 5.76	970.29 ± 9.52	0.57 ± 0.41	2.45 ± 0.79	3352.42 ± 2812.16	1.67 ± 0.79	87.88 ± 34.00	46.47 ± 17.77
15	1.72 ± 1.95	975.84 ± 2.96	0.66 ± 0.22	2.10 ± 0.43	1273.30 ± 368.17	0.81 ± 0.55	60.71 ± 8.39	43.93 ± 13.56
16	2.06 ± 3.61	971.10 ± 7.02	0.44 ± 0.24	3.13 ± 0.72	1709.53 ± 1409.05	3.04 ± 0.90	81.55 ± 17.15	35.07 ± 7.53
17	1.84 ± 2.03	979.06 ± 5.32	0.81 ± 0.50	2.30 ± 0.62	1424.75 ± 511.79	1.66 ± 1.21	70.48 ± 12.29	33.74 ± 15.37
18	2.10 ± 2.48	975.57 ± 6.90	0.25 ± 0.18	2.45 ± 0.73	1588.76 ± 601.41	2.35 ± 0.64	77.78 ± 11.55	42.89 ± 6.47
19	2.04 ± 1.74	977.05 ± 4.92	0.41 ± 0.10	2.61 ± 0.42	1529.41 ± 345.91	2.20 ± 0.70	74.42 ± 8.39	39.98 ± 8.06
20	2.92 ± 5.12	986.66 ± 3.57	0.42 ± 0.28	2.32 ± 0.51	1695.33 ± 1384.94	2.02 ± 1.22	74.11 ± 19.67	41.48 ± 14.24
21	2.88 ± 4.31	989.58 ± 4.84	0.46 ± 0.18	2.79 ± 0.61	2818.88 ± 1768.66	2.25 ± 1.05	79.79 ± 12.19	36.91 ± 8.69
22	2.06 ± 3.61	971.10 ± 7.02	0.44 ± 0.24	3.13 ± 0.72	1709.53 ± 1409.05	3.04 ± 0.90	81.55 ± 17.15	35.07 ± 7.53
23	1.56 ± 3.61	996.97 ± 9.77	0.68 ± 0.33	2.48 ± 0.40	1482.75 ± 795.01	2.28 ± 0.89	65.89 ± 7.96	57.83 ± 13.12
24	2.45 ± 3.79	974.08 ± 12.15	1.22 ± 0.70	1.84 ± 0.71	1700.32 ± 1262.69	1.77 ± 0.67	64.32 ± 19.65	48.49 ± 14.98
25	2.05 ± 2.99	982.02 ± 5.75	0.51 ± 0.23	2.27 ± 0.57	1435.53 ± 974.68	3.36 ± 1.16	74.26 ± 20.30	33.59 ± 10.83
26	2.02 ± 4.77	987.71 ± 4.82	0.46 ± 0.18	1.90 ± 0.59	1117.48 ± 558.34	1.76 ± 0.82	64.55 ± 13.50	58.65 ± 15.46
27	1.87 ± 1.71	976.22 ± 10.75	0.39 ± 0.09	3.33 ± 0.52	1559.86 ± 700.75	2.90 ± 0.76	77.65 ± 10.13	39.60 ± 6.14
28	2.80 ± 3.62	993.30 ± 10.95	0.36 ± 0.11	2.78 ± 0.52	1728.74 ± 851.36	1.86 ± 1.48	81.41 ± 11.36	40.69 ± 11.14
All	2.13 ± 3.16	982.19 ± 11.70	0.54 ± 0.39	2.57 ± 0.76	1663.64 ± 1113.03	2.14 ± 1.02	74.84 ± 16.53	41.98 ± 13.31

Table 2 (continue).

Subject No.	RESP	Skin Temperature	Steps	Activity Class					
	[bpm]	[C]		undefined	resting	other	biking	running	walking
1	16.70 ± 2.84	33.67 ± 1.38	8.29 ± 18.61	0.88	0.12	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	18.30 ± 4.28	34.10 ±1.12	5.56 ± 16.90	0.78	0.20	NaN	0.02	NaN	NaN
4	20.22 ± 4.93	35.15 ±1.09	3.55 ± 14.07	0.59	0.41	NaN	0.00	NaN	0.00
5	17.92 ± 6.78	34.86 ±0.93	4.59 ± 18.85	0.59	0.40	NaN	0.00	NaN	0.00
6	16.99 ± 2.72	33.50 ±1.79	6.96 ± 18.88	0.90	0.10	NaN	NaN	NaN	NaN
7	16.57 ± 4.07	34.48 ±1.39	3.90 ± 15.21	0.63	0.37	NaN	0.00	NaN	NaN
8	16.58 ± 3.29	34.44 ±1.48	3.25 ± 12.29	0.66	0.34	NaN	NaN	NaN	NaN
9	16.99 ± 4.42	34.51 ±1.09	4.88 ± 18.61	0.65	0.34	NaN	NaN	NaN	0.00
10	15.41 ± 2.33	35.62 ± 1.22	3.44 ± 13.72	0.74	0.26	NaN	NaN	NaN	NaN
11	16.80 ± 3.00	33.75 ± 1.10	3.88 ± 15.38	0.71	0.29	NaN	NaN	NaN	NaN
12	17.78 ± 2.79	34.62 ± 1.50	1.98 ± 10.10	0.85	0.15	NaN	NaN	NaN	NaN
13	19.75 ± 5.36	33.37 ± 1.22	6.73 ± 17.86	0.63	0.30	NaN	0.08	NaN	NaN
14	18.92 ± 6.37	33.56 ± 1.88	7.52 ± 17.50	0.70	0.17	NaN	0.14	NaN	NaN
15	14.53 ± 2.56	33.92 ± 1.00	2.56 ± 11.84	0.80	0.20	NaN	NaN	NaN	NaN
16	17.66 ± 4.19	33.87 ± 1.31	7.64 ± 28.07	0.71	0.27	NaN	0.00	0.01	NaN
17	15.94 ± 2.75	34.35 ± 1.64	5.72 ± 15.97	0.62	0.38	NaN	NaN	NaN	NaN
18	15.83 ± 3.70	33.95 ± 1.52	5.21 ± 16.38	0.77	0.23	NaN	NaN	NaN	NaN
19	17.61 ± 3.34	34.45 ± 1.26	4.92 ± 16.15	0.84	0.15	NaN	NaN	NaN	0.00
20	15.44 ± 4.31	33.99 ± 1.92	11.19 ± 32.45	0.52	0.44	NaN	NaN	0.04	NaN
21	17.91 ± 4.16	34.47 ± 1.13	2.44 ± 9.72	0.76	0.19	NaN	0.04	NaN	NaN
22	17.66 ± 4.19	33.87 ± 1.31	7.64 ± 28.07	0.71	0.27	NaN	0.00	0.01	NaN
23	16.89 ± 1.88	33.96 ± 1.43	1.70 ± 7.12	0.44	0.26	0.22	0.06	NaN	0.02
24	14.91 ± 2.51	32.32 ± 1.96	13.29 ± 32.38	0.19	0.54	0.07	0.07	0.13	0.01
25	15.21 ± 2.44	33.84 ± 1.35	5.02 ± 14.97	0.15	0.58	0.1	0.08	0.00	0.09
26	14.88 ± 2.05	35.23 ± 0.91	3.94 ± 13.53	0.12	0.63	0.14	0.03	NaN	0.08
27	16.75 ± 1.86	34.87 ± 1.28	6.81 ± 19.46	0.52	0.23	0.15	0.05	0.02	0.03
28	19.25 ± 4.02	35.22 ± 1.07	6.02 ± 17.73	0.50	0.20	0.16	0.07	0.02	0.06
All	16.97 ± 4.05	34.23 ± 1.54	5.35 ± 18.05	0.63	0.29	0.03	0.02	0.01	0.01

Table 3: Subject-level mean ± standard deviation (std) of the 11 numerical sensor parameters and category ratios for the categorical parameter 'ActivityClass' over the 405 selected recorded days. HR: heart rate; HRV: heart rate variability; RESP: respiratory rate; mbar: millibar; (m)s: (milli)seconds; kOhm: kilo Ohm; bpm: beats per minute; C: Celsius degree.

Subject No.	Activity Counts	Barometer [mbar]	Blood Perfusion	Blood Pulse Wave	Energy Expenditure [calories/s]	Galvanic Skin Response [kOhm]	HR [bpm]	HRV [ms]
1	3.11 ± 6.04	989.45 ± 9.99	0.76 ± 0.38	2.31 ± 0.61	1678.65 ± 1278.15	2.46 ± 1.24	71.55 ± 16.95	44.4 ± 11.78
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	1.91 ± 3.08	986.61 ± 7.72	0.52 ± 0.29	2.63 ± 0.94	1775.00 ± 1232.43	2.04 ± 1.19	80.01 ± 20.64	36.94 ± 9.69
4	1.77 ± 2.93	978.45 ± 31.73	0.63 ± 0.48	3.74 ± 1.04	1750.34 ± 1206.61	2.47 ± 1.20	90.87 ± 15.81	30.87 ± 11.30
5	2.11 ± 2.76	983.91 ± 12.77	0.31 ± 0.21	3.02 ± 0.85	1849.40 ± 1157.33	2.19 ± 1.15	87.36 ± 17.78	45.48 ± 14.34
6	1.97 ± 2.02	984.94 ± 10.11	0.39 ± 0.16	2.51 ± 0.48	1541.59 ± 512.41	2.52 ± 0.38	74.16 ± 10.03	39.95 ± 10.69
7	1.69 ± 3.23	982.41 ± 6.03	0.39 ± 0.22	2.42 ± 0.50	1417.45 ± 611.80	2.07 ± 0.99	70.83 ± 11.99	43.32 ± 12.27
8	1.84 ± 2.74	975.5 ± 22.84	0.46 ± 0.29	2.36 ± 0.62	1614.72 ± 880.06	2.06 ± 0.70	70.46 ± 12.04	42.22 ± 12.56
9	2.36 ± 5.14	982.62 ± 8.97	0.53 ± 0.28	2.01 ± 0.63	1583.24 ± 1019.05	2.44 ± 1.03	64.87 ± 17.78	32.64 ± 10.61
10	1.70 ± 2.71	981.80 ± 7.75	0.30 ± 0.19	2.59 ± 0.50	1542.20 ± 802.46	2.72 ± 1.01	72.73 ± 13.54	39.03 ± 10.09
11	1.74 ± 2.83	987.43 ± 8.23	0.60 ± 0.43	2.81 ± 0.58	1488.62 ± 640.04	2.49 ± 0.95	75.36 ± 12.07	35.21 ± 12.24
12	1.86 ± 3.09	986.19 ± 10.03	0.55 ± 0.31	2.70 ± 0.51	1600.05 ± 925.46	2.36 ± 0.97	72.53 ± 13.17	42.62 ± 12.47
13	1.87 ± 2.58	974.75 ± 25.45	0.57 ± 0.39	2.99 ± 0.78	1651.58 ± 964.85	2.31 ± 1.17	77.82 ± 14.64	32.55 ± 10.49
14	2.27 ± 3.87	982.63 ± 7.47	0.40 ± 0.27	2.51 ± 0.61	1886.06 ± 1637.74	2.46 ± 1.17	75.23 ± 20.24	37.41 ± 12.22
15	1.88 ± 4.00	975.11 ± 20.34	0.53 ± 0.28	2.25 ± 0.50	1503.05 ± 858.32	2.61 ± 1.33	66.06 ± 17.18	40.80 ± 12.96
16	2.20 ± 4.13	975.36 ± 12.45	0.42 ± 0.23	2.96 ± 0.72	1641.39 ± 1082.66	2.64 ± 1.18	80.19 ± 16.17	38.40 ± 10.47
17	2.40 ± 5.48	980.46 ± 9.81	0.39 ± 0.29	2.56 ± 0.63	1566.64 ± 960.28	2.33 ± 1.27	74.17 ± 14.50	29.26 ± 12.59
18	1.87 ± 2.61	974.76 ± 8.76	0.17 ± 0.15	2.30 ± 0.67	1450.82 ± 574.05	2.15 ± 0.83	73.07 ± 11.51	52.56 ± 19.90
19	2.12 ± 3.48	995.25 ± 23.24	0.37 ± 0.20	2.38 ± 0.51	1482.87 ± 571.87	2.59 ± 1.17	70.27 ± 12.64	40.81 ± 10.01
20	2.24 ± 3.82	985.25 ± 7.39	0.37 ± 0.31	2.24 ± 0.69	1593.32 ± 1029.75	1.29 ± 1.03	70.47 ± 18.82	39.81 ± 13.68
21	2.31 ± 3.87	982.86 ± 5.39	0.44 ± 0.17	2.68 ± 0.52	1745.57 ± 1084.40	1.85 ± 1.17	73.24 ± 12.48	34.94 ± 10.15
22	2.20 ± 4.13	975.36 ± 12.45	0.42 ± 0.23	2.96 ± 0.72	1641.39 ± 1082.66	2.64 ± 1.18	80.19 ± 16.17	38.40 ± 10.47
23	1.81 ± 2.94	985.15 ± 12.48	0.63 ± 0.39	2.46 ± 0.59	1734.92 ± 1315.03	1.87 ± 0.89	71.65 ± 14.51	46.46 ± 17.01
24	2.45 ± 4.11	981.70 ± 20.33	0.66 ± 0.45	2.05 ± 0.72	2574.10 ± 2148.21	1.98 ± 1.03	67.74 ± 19.07	48.87 ± 17.40
25	2.25 ± 3.82	980.52 ± 9.30	0.51 ± 0.29	2.28 ± 0.57	2547.26 ± 1740.22	2.53 ± 0.99	72.76 ± 15.88	37.36 ± 12.74
26	1.89 ± 3.38	985.48 ± 9.03	0.52 ± 0.27	2.03 ± 0.57	1783.12 ± 1119.86	2.22 ± 1.17	64.92 ± 14.22	55.04 ± 17.16
27	1.92 ± 2.41	974.29 ± 11.56	0.41 ± 0.23	2.87 ± 0.68	1599.37 ± 770.10	2.29 ± 0.94	72.69 ± 11.78	46.23 ± 13.57
28	2.36 ± 3.60	987.96 ± 9.54	0.42 ± 0.15	2.49 ± 0.55	1738.73 ± 832.59	2.25 ± 1.40	75.39 ± 13.68	46.71 ± 18.91
All	2.10 ± 3.62	982.56 ± 16.01	0.54 ± 0.37	2.38 ± 0.74	1927.57 ± 1499.54	2.15 ± 1.08	70.95 ± 16.61	45.10 ± 16.56

Table 3 (continue).

Subject No.	RESP	Skin Temperature	Steps	Activity Class					
	[bpm]	[C]		undefined	resting	other	biking	running	walking
1	16.44 ± 3.85	34.48 ± 1.25	8.74 ± 26.44	0.61	0.37	NaN	0.00	0.01	0.00
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	18.62 ± 5.00	34.21 ± 1.53	2.84 ± 13.10	0.5	0.44	NaN	0.06	NaN	0.00
4	21.63(± 4.84	33.77 ± 1.98	4.11 ± 17.32	0.39	0.59	NaN	0.02	NaN	0.00
5	20.40(± 6.10	33.96 ± 2.91	7.68 ± 24.07	0.64	0.3	NaN	0.05	NaN	0.01
6	17.72(± 3.25	33.65 ± 1.42	6.67 ± 18.53	0.89	0.11	NaN	0.01	NaN	0.00
7	16.45(± 3.75	34.48 ± 1.19	3.65 ± 15.48	0.47	0.53	NaN	0.00	NaN	NaN
8	16.79(± 3.97	33.61 ± 2.07	3.91 ± 15.18	0.51	0.49	NaN	NaN	NaN	0.00
9	15.36(± 4.21	33.34 ± 1.76	9.03 ± 29.67	0.41	0.55	NaN	0.00	0.01	0.03
10	15.90(± 4.28	35.46 ± 2.15	4.41 ± 19.00	0.4	0.59	NaN	0.01	0.00	0.00
11	17.60(± 4.09	33.96 ± 1.39	2.72 ± 12.69	0.46	0.53	NaN	0.01	NaN	NaN
12	16.25(± 3.89	33.50 ± 2.23	4.35 ± 19.89	0.52	0.48	NaN	0.00	0.00	0.00
13	18.41(± 5.30	33.79 ± 1.50	3.80 ± 14.55	0.44	0.53	NaN	0.03	NaN	0.00
14	17.07(± 4.75	34.66 ± 1.37	2.76 ± 12.28	0.48	0.48	NaN	0.04	NaN	0.00
15	16.93(± 6.28	33.06 ± 3.41	3.00 ± 15.49	0.46	0.53	NaN	NaN	0.00	0.00
16	18.62(± 4.37	33.92 ± 2.21	6.49 ± 24.30	0.51	0.48	NaN	0.00	0.01	0.00
17	17.37(± 4.10	34.78 ± 1.52	3.97 ± 16.88	0.38	0.6	NaN	0.02	0.00	0.00
18	14.73(± 3.96	35.33 ± 1.55	2.69 ± 12.13	0.49	0.51	NaN	0.00	NaN	NaN
19	16.00(± 4.33	34.70 ± 1.40	5.37 ± 18.59	0.62	0.38	NaN	NaN	NaN	0.00
20	14.87(± 4.00	34.52 ± 3.03	7.20 ± 25.56	0.43	0.55	NaN	NaN	0.02	0.00
21	17.17(± 3.59	34.48 ± 1.52	3.47 ± 13.62	0.5	0.48	NaN	0.02	NaN	0.00
22	18.62(± 4.37	33.92 ± 2.21	6.49 ± 24.30	0.51	0.48	NaN	0.00	0.01	0.00
23	17.22(± 2.75	34.26 ± 2.11	4.27 ± 17.33	0.15	0.54	0.19	0.06	0.02	0.03
24	14.98(± 3.57	33.44 ± 2.20	7.80 ± 23.64	0.2	0.49	0.13	0.07	0.06	0.05
25	16.65(± 3.52	33.98 ± 1.74	5.25 ± 16.03	0.31	0.43	0.12	0.07	0.01	0.06
26	15.73(± 2.81	35.29 ± 1.26	4.19 ± 15.52	0.14	0.57	0.19	0.03	0.01	0.05
27	16.73(± 3.34	34.63 ± 1.44	4.25 ± 16.19	0.48	0.37	0.09	0.04	0.01	0.02
28	17.67(± 4.05	34.89 ± 1.21	4.12 ± 14.22	0.3	0.46	0.13	0.06	0.01	0.04
All	16.43(± 3.95	34.16 ± 2.06	5.28 ± 19.23	0.31	0.49	0.1	0.04	0.02	0.03

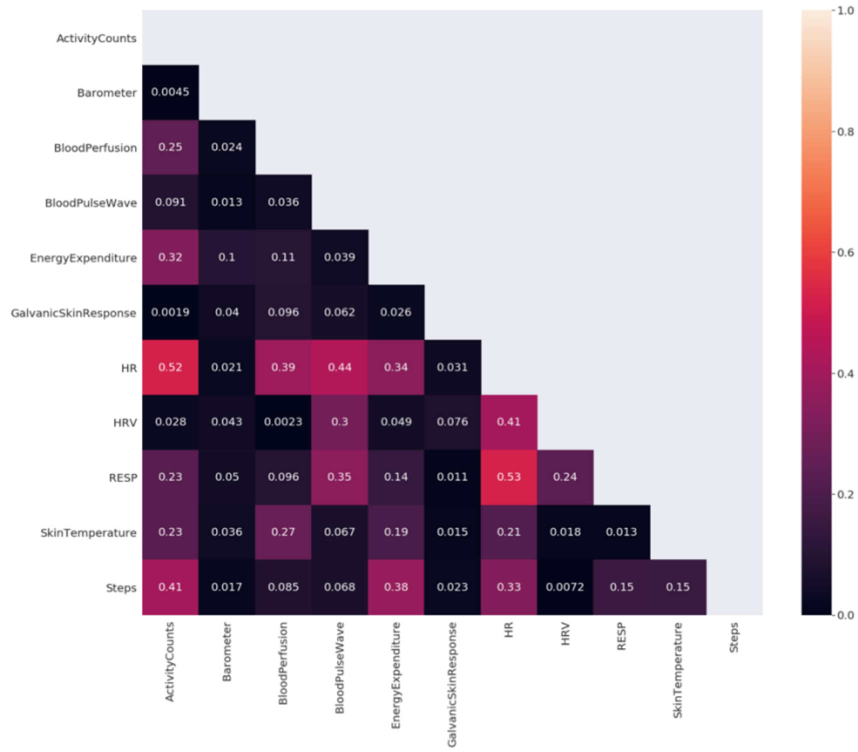

Figure 5: Absolute Spearmen correlation matrix of the 11 numerical sensor parameters at 1-minute temporal resolution.

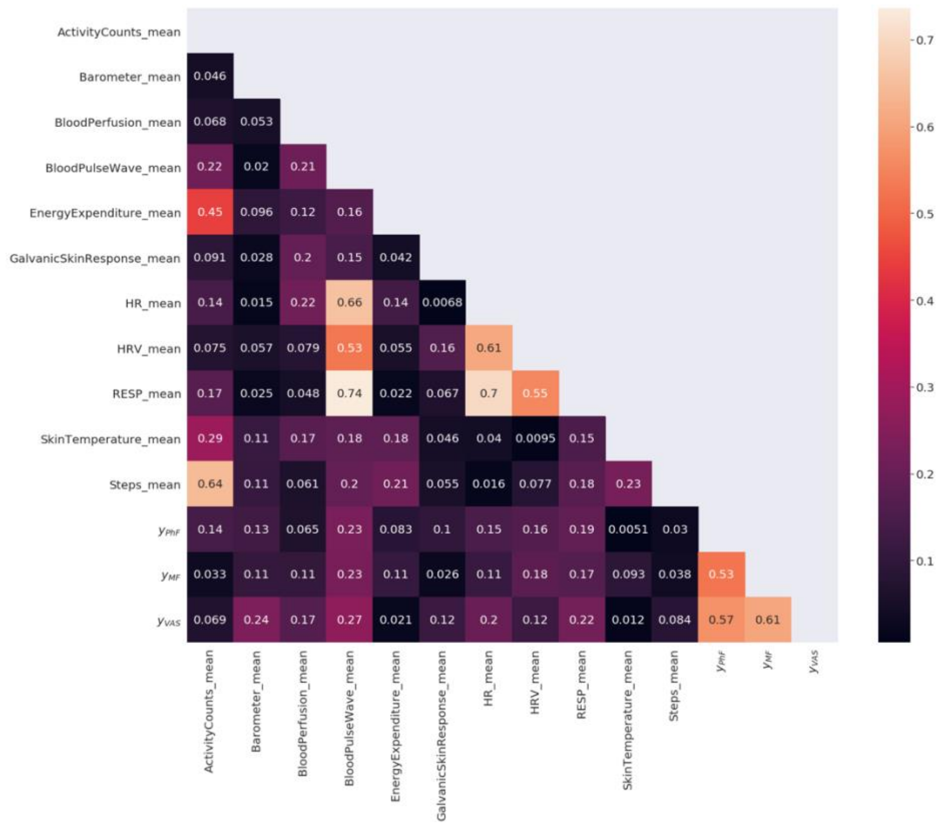

Figure 6: Absolute Spearmen correlation matrix of the daily average of the 11 numeric sensor parameters and daily fatigue labels.

Statistical significance test:

t-test with independent samples and Bonferroni correction. p-value annotation (in Figs.3-5):

ns: 5.00e-02 < p ⩽ 1.00e+00;

∗ : 1.00e−02 < p ⩽ 5.00e−02;

∗∗ : 1.00e−03 < p ⩽ 1.00e − 02;

∗ ∗ ∗ : 1.00e – 04 < p ⩽ 1.00e − 03;

∗ ∗ ∗∗ : p ⩽ 1.00e – 04.

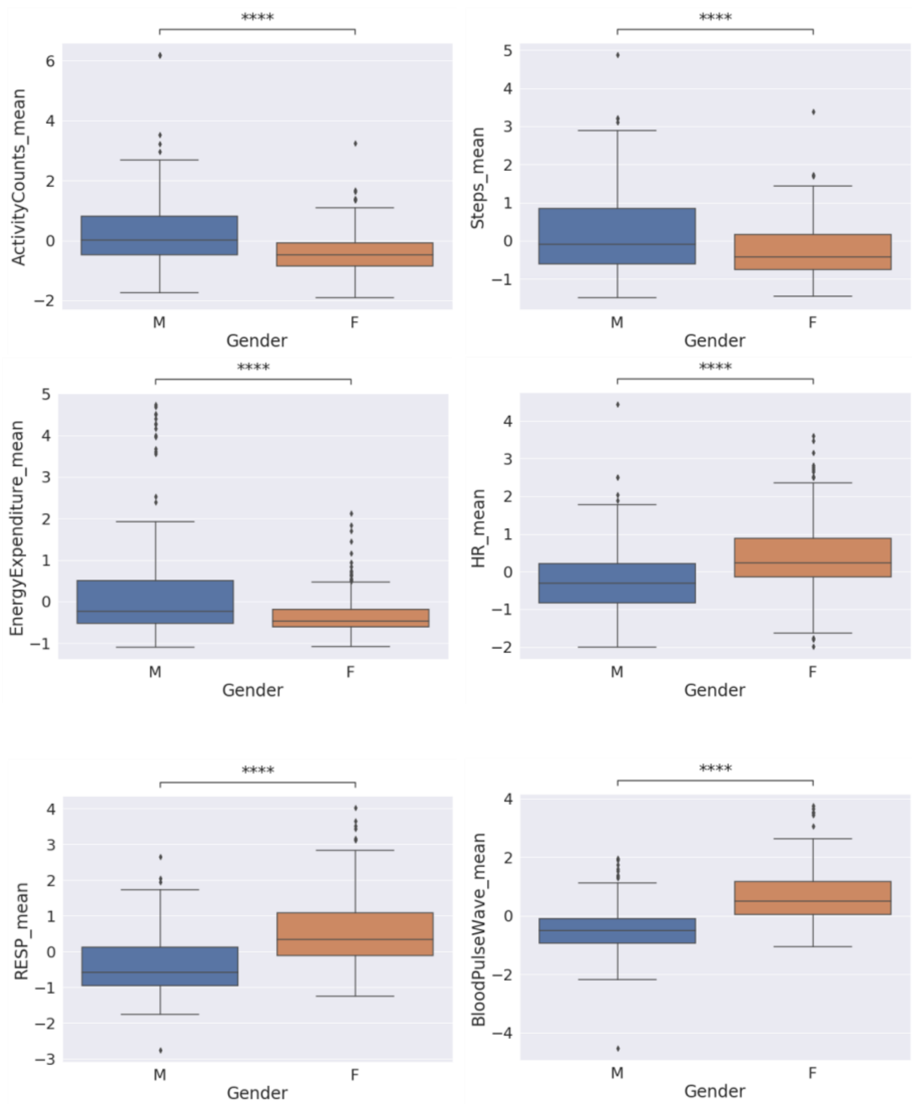

Figure 7: Selected boxplots and statistical significance differences of sensor features for different genders.

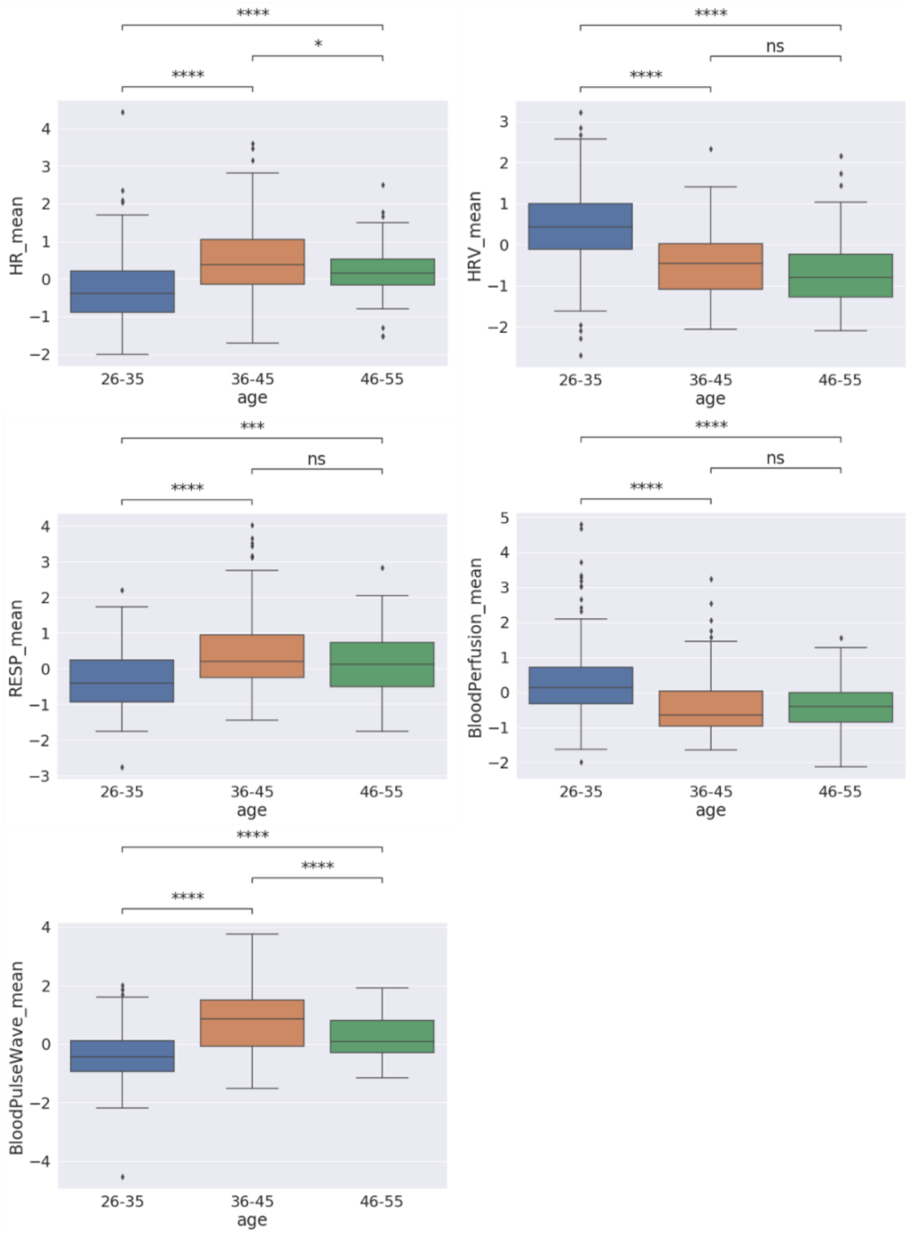

Figure 8: Selected boxplots and statistical significance differences of sensor features for different age-range groups.

A2. Missing Data Imputation

Method

Missing sensor data are ubiquitous in the real world due to uncontrollable factors linked to the subjects wearing these devices. Yet having complete time series is important to reduce noise when computing features and to ensure fixed-length input for DL models. To reduce the negative effect of information loss on data quality and hence downstream tasks due to missing time series segments in the sensor data, we imputed incomplete data. Prior to imputing missing data, we normalized the sensor data by using Z-score for each parameter independently and considering all subjects in one distribution. Then, we imputed missing data in a two-step approach solely based on sensor data, without using patient-reported fatigue labels. Firstly, we assumed that parameters undergo small changes within a short time window. Chosen experimentally, where data gaps in a sensor parameter were ≤ 2 minutes and where at least 5 minutes of data before or after the gap were recorded, the most adjacent sample was copied. This forward and backward filling step for short time windows is commonly used in time series data [5, 6]. After that, for imputing longer missing data segments in each parameter independently, we benchmarked:

- Global mean
- Subject-specific periodic mean over a 6-hour window
- K-nearest neighbors (KNN) with K = 10
- A recently developed set of recurrent neural networks (RNNs) [7]
    - Unidirectional correlated recurrent imputation (RITS)
    - Unidirectional uncorrelated recurrent imputation (RITS-I)
    - Bidirectional correlated recurrent imputation (BRITS)
    - Bidirectional uncorrelated recurrent imputation (BRITS-I).

These state-of-the-art RNN-based algorithms treated missing values as variables in the RNN graph, which are updated through backpropagation and hence directly learnt from recurrent dynamics without specific assumptions of underlying data distributions [7].

Implementation details of RNN-based imputation models (RITS, RITS-I, BRITS, BRITS-I)

Library: pytorch

Hyperparameters:

- Batch size = 32
- Learning rate = 0.001
- Dimension of hidden recurrent state = 64

Evaluation Criteria

In order to evaluate the imputation results of all the compared imputation models, we manually masked 10% of non-missing data as NaN. Based on the pattern of missing time windows observed on our data, the generated masking was split as follows:

- 50% were masked out as single, randomly selected points
- 20% in 1-hour window of consecutive points
- 20% in 2-hour windows
- 10% of 6-hour windows.

For each parameter, evaluation metrics were mean absolute error (MAE) and mean relative error (MRE), computed on masked data points: MAE = $\frac{\sum_{i=1}^{n}|\hat{x}_i - x_i|}{n}$ and MRE = $\frac{\sum_{i=1}^{n}|\hat{x}_i - x_i|}{\sum_{i=1}^{n}|x_i|}$, with x_i the ground truth value for the masked time point *i*, $\hat{x}_i$ its corresponding predicted value from the imputation algorithm, and *n* the total number of masked points for all subjects in one distribution.

Results

On average, for 45% of the observations, the length of missing data windows was ≤ 2 hours; yet for 12% of the cases, it was > 12 hours, see details in Figure 9. Table 4 summarizes the results of the benchmarked imputation models on the masked validation set. The best performing method was RITS-I with MAE (MRE) of 0.3 (40.8%), while the worst results are obtained when using global mean. Figure 10 shows the MAE and MRE curves, with increasing number of epochs for the four models and the stop of training at the minimal loss. Errors decrease smoothly for all the models with RITS-I achieving faster the lowest MAE and MRE. Qualitative evaluation of our results on actual missing data suggests that the imputed values are in a realistic range for the best performing methods (RITS-I and KNN), as shown in a representative example in Figure 11. We used data generated from RITS-I to impute the corresponding missing values for all downstream analyses presented in the main article.

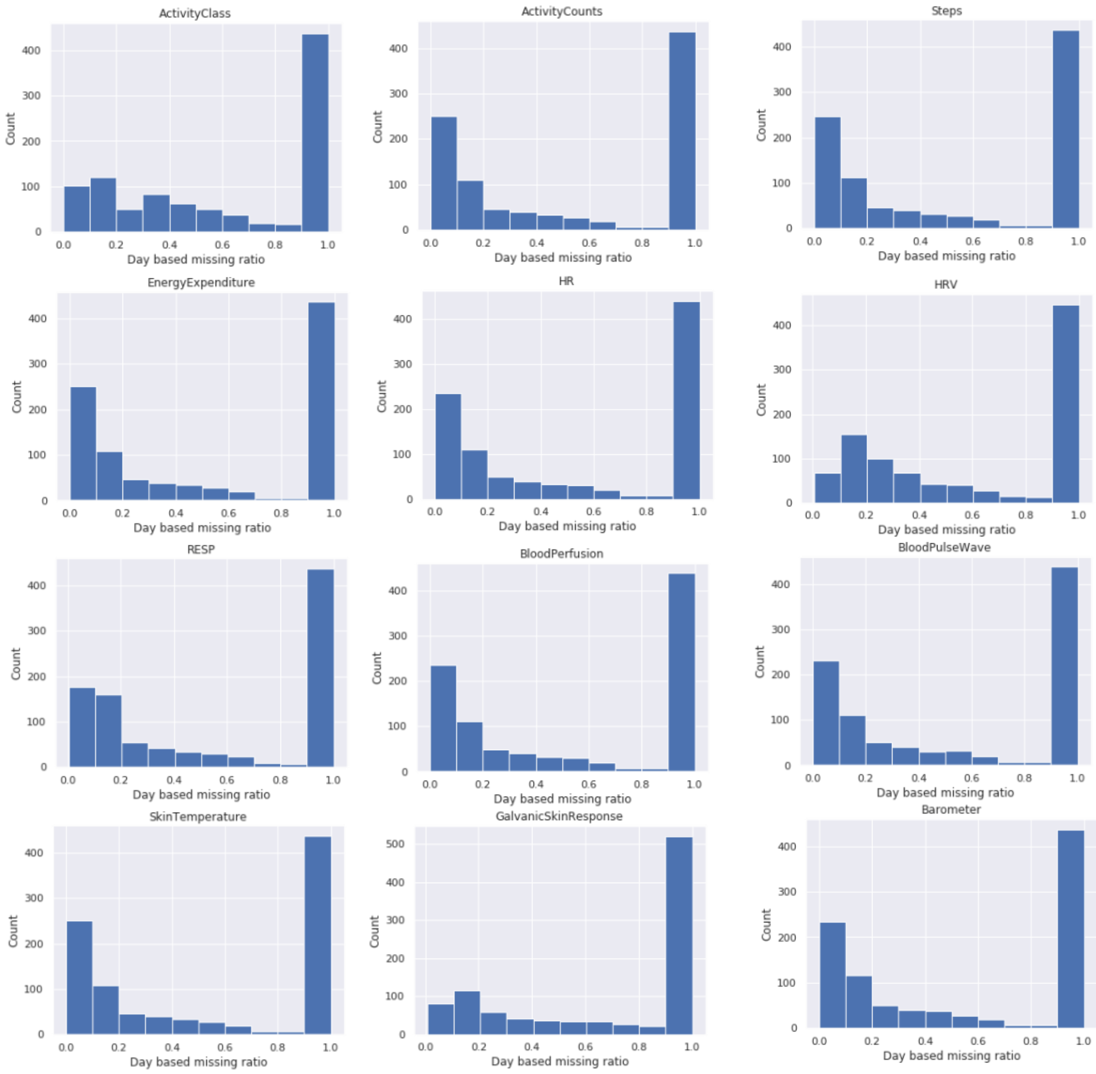

Figure 9: Distribution of daily-based missing rated for all sensor parameters and all subjects.

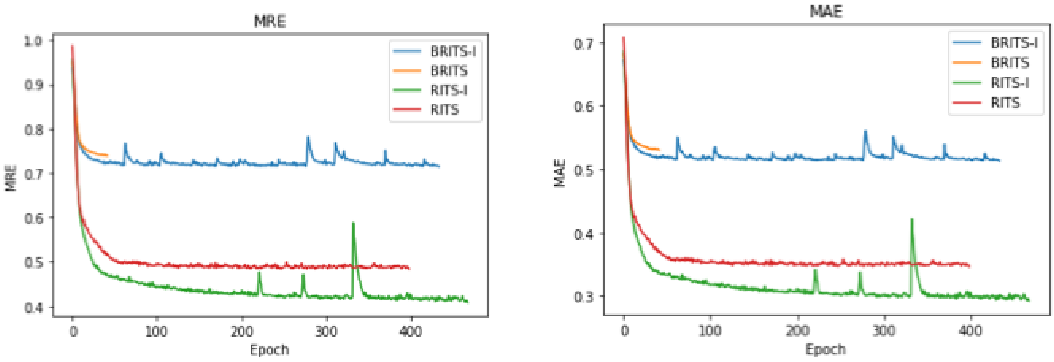

Figure 10: Mean relative error (MRE, left) and mean absolute error (MAE, right) training curves of recurring imputation networks.

Table 4. Missing data imputation results. MAE: mean absolute error. MRE: mean relative error. Best results are highlighted in bold font.

Imputation method	**MAE (MRE%)**
Mean	0.71 (100%)
Periodic mean within subject	0.61 (86.20%)
KNN	0.45 (63.47%)
RITS-I	**0.29 (40.80%)**
RITS	0.35 (48.42%)
BRITS-I	0.51 (71.61%)
BRITS	0.53 (73.91%)

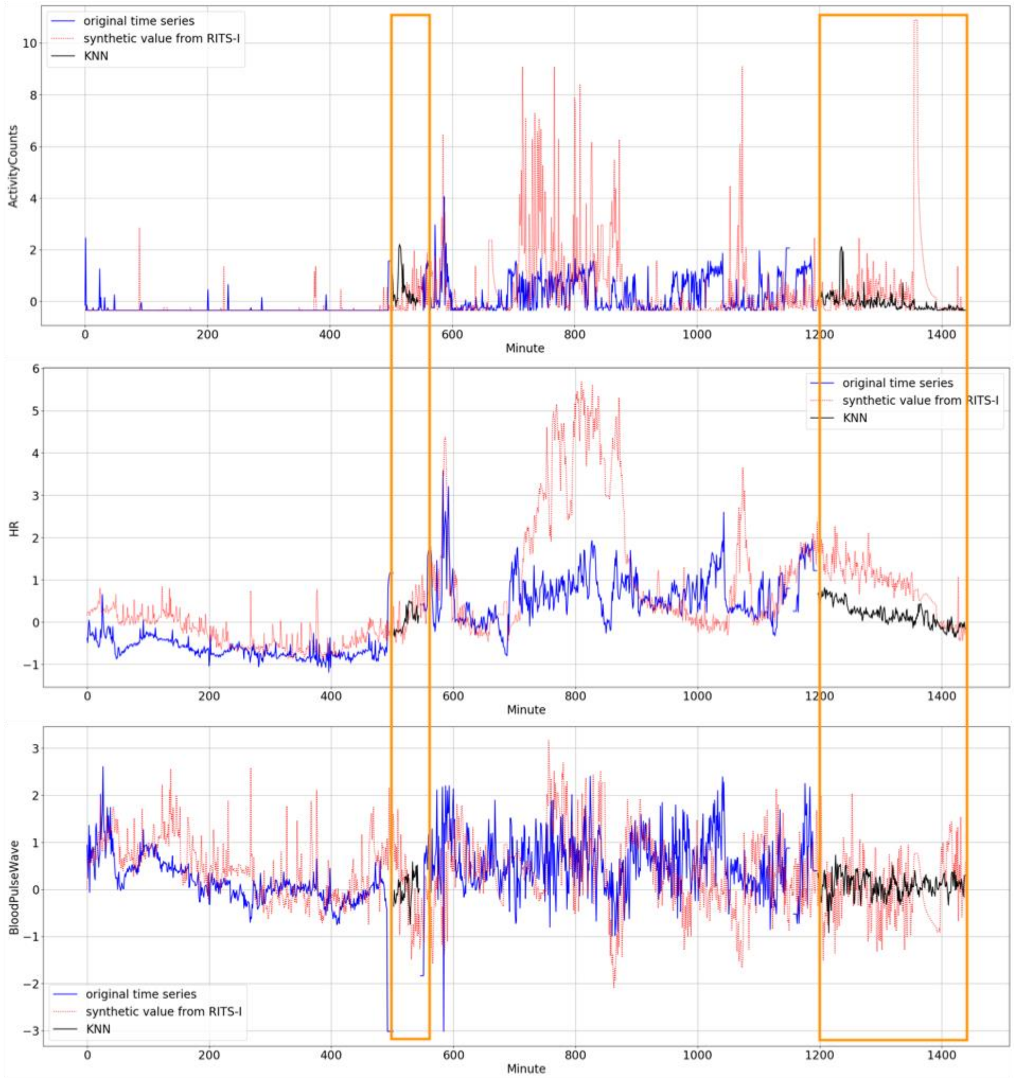

Figure 11: Example of original daily time series (blue) and imputed time series using the proposed RITS-I (red) vs. KNN (black) for ActivityCounts, HR and BloodPulsewave. Time windows of original missing values are framed by yellow boxes.

A3. Implementation Details of Causal CNN

Causal Convolutional Neural Network (cCNN) [8]:

Library: pytorch

Hyperparameters:

- Number of hidden neurons in the 1st 1-D Conv ∈ {64;8}
- Number of hidden neurons in the 2rd 1-D Conv ∈ {64;8}
- Kernel size of the 1st 1-D Conv ∈ {5;3}
- Kernel size of the 2rd 1-D Conv ∈ {5;3}
- Number of hidden neurons in the 1st fully connected layer ∈ {128;64}
- Number of hidden neurons in the 2rd fully connected layer ∈ {128;64}
- Learning rate ∈ {0.000001;0.0001;0.001}
- Batch size = 4
- L2 regularizer ∈ {0.0001;0.001}
- Activation function = *ReLu*
- Drop out rate at 1st and 2rd fully connected layer ∈ {0.7;0.5}
- Random selection of negative samples in the triplet loss function

The model was trained in an unsupervised way using 5-fold cross validation (CV).

Model complexity: 84660 parameters.

A4. Implementation Details of Classification Models

Logistic regression:

Library: scikit-learn

Hyperparameters:

- Penalty ∈ {L1, L2}
- *Inverse of regularization strength C* ∈ {0, 4, 10}

Model complexity: 15 hyperparameters.

Support Vector Machine (SVM):

Library: scikit-learn

Hyperparameters:

- Kernel type ∈ {linear; radial basis function (rbf); 3rd degree polynomial (poly)}
- L2 regularization parameter *C* ∈ {0.001; 0.01; 0.1; 10; 100; 1000}
- Kernel coefficient, *γ* ∈ {0.01; 0.1; 1} (for rbf and poly kernel types)

Model complexity: 15 hyperparameters.

Random Forest (RF):

Library: scikit-learn

Hyperparameters:

- Number of trees, N_estimators ∈ {3; 10; 50; 100; 300; 1000}
- The number of features to consider when looking for the best split, Max features ∈ {auto; sqrt; log2}
- The maximum depth of the tree, Max depth ∈ {10; 30; 50; 100}

- The function to measure the quality of a split, Criterion ∈ {gini; entropy}
- The minimum number of samples required to split an internal node, Min samples split ∈ {2; 5; 10}

Model complexity: 17 hyperparameters.

1-D CNN:

The 1-D CNN is composed by two 1-D convolution layers followed by two fully connected layers and an output softmax layer.

Library: pytorch

Hyperparameters:

- Number of hidden neurons in the 1st 1-D Conv ∈ {8; 64}
- Number of hidden neurons in the 2rd 1-D Conv ∈ {8; 64}
- Kernel size of the 1st 1-D Conv ∈ {3; 5}
- Kernel size of the 2rd 1-D Conv ∈ {3; 5}
- Number of hidden neurons in the 1st fully connected layer ∈ {64; 128}
- Number of hidden neurons in the 2rd fully connected layer ∈ {64; 128}
- Learning rate ∈ {0.000001; 0.0001; 0.001}
- Batch size = 4
- L2 regularizer ∈ {0.0001; 0.001}
- Activation function = *ReLu*
- Drop out rate at 1st and 2rd fully connected layer ∈ {0.5; 0.7}

Model complexity: 2960908 parameters

Model Training and Evaluation Criteria

Five-fold nested cross-validation with train-validation-test data split was used to select model hyperparameters (inner-loop) and train model parameters (outer loop), see Figure 12. Considering the unbalanced label class, both inner and outer corss-validation followed Stratified K-fold cross validation iterator provided by scikit-learn [9], which will keep the split fold with the same percentage of sample for each class. We gave priority to stratified label classes instead of patients due to the size and unbalanced nature of the dataset. In future work, we will use patient stratification to train larger and more balanced datasets. We normalized (Z-score) features only in the training set and used the resulting mean and std to normalize the test set. To mitigate for unbalanced labels, we added class weights (ratio between number of samples from each class and the total number of samples in the training set) to the loss function of the prediction models [10]. Furthermore, we optimized all models based on weighted F1 score.

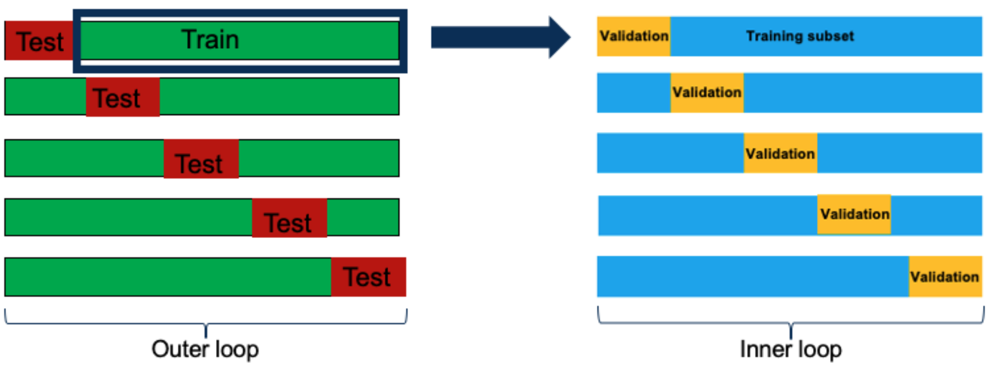

Figure 12: 5 folds nested cross-validation.

A5. Classification Results

Table 5: Prediction results of binary labels of physical and mental fatigue. Evaluation metrics are summarized as cross-validation mean ± standard deviation. Best results are highlighted in bold font.

	Model	Accuracy [%]	Weighted precision	Weighted recall	Weighted F1 score
Physical fatigue	Logistic regression	70.86 ± 0.60	0.50 ± 0.01	0.71 ± 0.01	0.59 ± 0.01
	SVM	53.58 ± 20.56	0.33 ± 0.20	0.54 ± 0.21	0.40 ± 0.22
	RF	**71.85 ± 1.44**	0.70 ± 0.05	0.72 ± 0.01	0.65 ± 0.02
	1D-CNN	58.02 ± 13.37	0.53 ± 0.05	0.58 ± 0.13	0.52 ± 0.07
	cCNN+RF	71.40 ± 3.92	0.70 ± 0.03	0.73 ± 0.03	**0.71 ± 0.04**
Mental fatigue	Logistic regression	64.44 ± 0.01	0.42 ± 0.01	0.64 ± 0.01	0.51 ± 0.01
	SVM	46.42 ± 14.51	0.24 ± 0.14	0.46 ± 0.15	0.31 ± 0.16
	RF	64.69 ± 3.37	0.63 ± 0.04	0.65 ± 0.03	0.62 ± 0.02
	1D-CNN	43.95 ± 8.64	0.38 ± 0.13	0.44 ± 0.09	0.37 ± 0.11
	cCNN+RF	**66.20 ± 4.58**	0.65 ± 0.05	0.66 ± 0.05	**0.65 ± 0.05**

Table 6: Pair-wise p-value (McNemar's test) of prediction accuracy of physical fatigue for all the models. Statistical significant pairs at 0.01 level are highlighted in bold font.

	Logistic regression	**SVM**	**RF**	**1D-CNN**
Logistic regression				
SVM	**4.5e-10**			
RF	**0.02**	**3.76e-06**		
1D-CNN	**4.3e-13**	0.22	**3.15e-04**	
cCNN+RF	0.24	**5.56e-7**	0.57	**6.43e-6**

Table 7: Pair-wise p-value (McNemar's test) of prediction accuracy of mental fatigue for all the models. Statistical significant pairs at 0.01 level are highlighted in bold font.

	Logistic regression	**SVM**	**RF**	**1D-CNN**
Logistic regression				
SVM	**3.28e-7**			
RF	0.56	**2.10e-06**		
1D-CNN	**5e-5**	0.63	**4.31e-08**	
cCNN+RF	0.81	**1.67e-08**	0.5	**1.17e-09**

*Table 8: The first 25 (10%) predictive features of RF for predicting **physical** fatigue. CV: cross-validation; HR: heart rate; HRV: heart rate variability; RESP: respiratory rate; PSDi: i-th peak of power spectral density.*

Features	**Mean importance of 5-fold CV**	**Standard deviation**
EnergyExpenditure_PSD2	0.037491	0.004525
EnergyExpenditure_sum	0.033137	0.001844
HRV_PSD1	0.031363	0.002698
EnergyExpenditure_PSD1	0.030417	0.003656
ActivityCounts_PSD1	0.029888	0.004496
RESP_PSD1	0.028896	0.003063
HR_PSD3	0.028837	0.002735
ActivityCounts_PSD2	0.028721	0.002845
EnergyExpenditure_max	0.028657	0.002305
Steps_sum	0.028155	0.001736
EnergyExpenditure_PSD3	0.027721	0.004864
HR_PSD1	0.026295	0.002599
HRV_PSD2	0.025966	0.00246
HR_PSD2	0.02572	0.002458
HR_PSD4	0.025341	0.003575
HRV_PSD4	0.025169	0.004519
EnergyExpenditure_percentile95	0.025059	0.002364
ActivityCounts_PSD3	0.024726	0.002819
EnergyExpenditure_PSD4	0.024488	0.003051
HRV_PSD3	0.023783	0.001795
RESP_PSD3	0.022936	0.002222
ActivityCounts_PSD4	0.022344	0.003022
RESP_PSD2	0.02073	0.002451
RESP_PSD4	0.019793	0.001696
ActivityCounts_sum	0.019714	0.001111

*Table 9: The first 25 (10%) predictive features of RF for predicting **mental** fatigue. CV: cross-validation; HR: heart rate; HRV: heart rate variability; RESP: respiratory rate; PSDi: i-th peak of power spectral density; ActivityClass_0: count of time intervals spent on undefined, non-rest physical activities.*

Features	**Mean importance of 5-fold CV**	**Standard deviation**
RESP_PSD1	0.03903	0.023689
HRV_PSD4	0.031783	0.006006
RESP_PSD2	0.029493	0.008804
ActivityCounts_PSD1	0.027972	0.003289
HR_PSD3	0.027466	0.004922
EnergyExpenditure_PSD1	0.027441	0.009318
Steps_sum	0.026952	0.003724
EnergyExpenditure_max	0.02664	0.006314

RESP_PSD4	0.026509	0.001777
HR_PSD1	0.025285	0.008394
RESP_PSD3	0.025222	0.004313
HR_PSD2	0.024628	0.007328
EnergyExpenditure_PSD4	0.024594	0.009093
EnergyExpenditure_PSD3	0.022599	0.009746
HRV_PSD1	0.022253	0.005766
ActivityClass_0	0.022178	0.002186
HR_PSD4	0.02158	0.003162
HRV_PSD2	0.021232	0.001718
ActivityCounts_sum	0.021053	0.006581
ActivityCounts_PSD3	0.020318	0.006187
ActivityCounts_PSD2	0.019848	0.003799
ActivityCounts_PSD4	0.019266	0.00479
EnergyExpenditure_percentile5	0.018877	0.00446
EnergyExpenditure_sum	0.018521	0.004222
EnergyExpenditure_PSD2	0.018198	0.006033

A6. Hierarchical clustering results

At height = 50 of the hierarchical clustering tree in the feature space (see horizontal red line in Fig. 3 in the main paper), we obtain three feature sets:

- FeatureGroup 1 mainly composed by physical activity features, e.g. from ActivityCounts and EnergyExpenditure, but also HRV and BloodPerfusion
- FeatureGroup 2 by vital sign parameters HR, RESP and BloodPulseWave
- FeatureGroup 3 by physiological, non-clinical measures, e.g. SkinTemperature, GalvanicSkinResponse and Barometer.

At height = 40 of the tree in the observation space (see vertical red line in Fig. 3 in the main paper), we obtain three clusters. Feature distributions in these groups are summarized in Figures 13-15 below. The mean values of sensor parameters in each group are listed in Table 2.

Statistical significance test:

t-test with independent samples and Bonferroni correction. p-value annotation (in Figs.3-5):

ns: 5.00e-02 < p ⩽ 1.00e+00;

∗ : 1.00e−02 < p ⩽ 5.00e−02;

∗∗ : 1.00e−03 < p ⩽ 1.00e − 02;

∗ ∗ ∗ : 1.00e – 04 < p ⩽ 1.00e − 03;

∗ ∗ ∗∗ : p ⩽ 1.00e – 04.

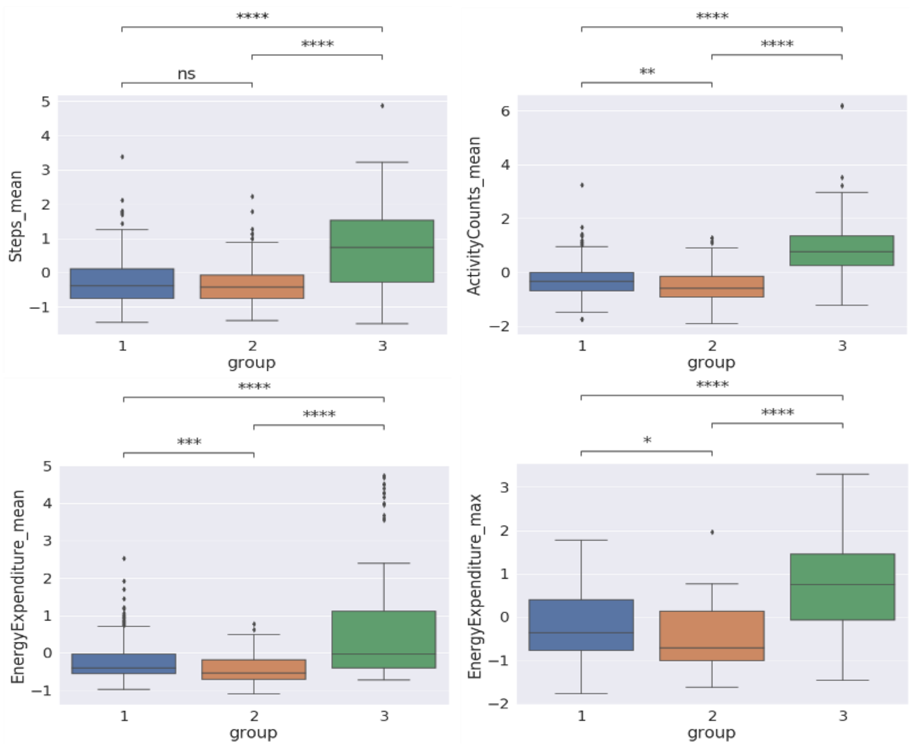

Figure 13: Selected boxplots and statistical significance differences of physical activity features for the three clustered groups.

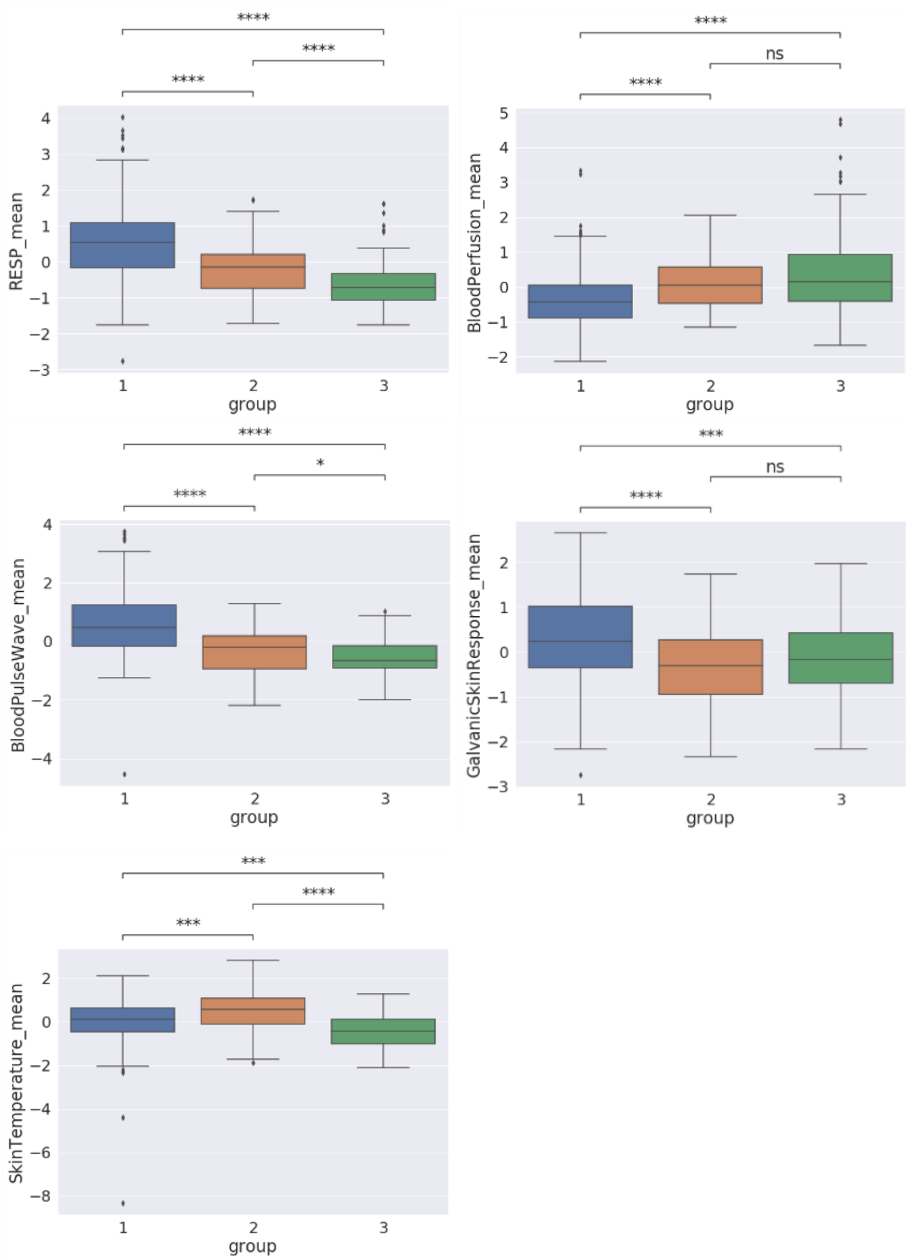

Figure 14: Selected boxplots and statistical significance differences of vital sign and physiological features for the three clustered groups.

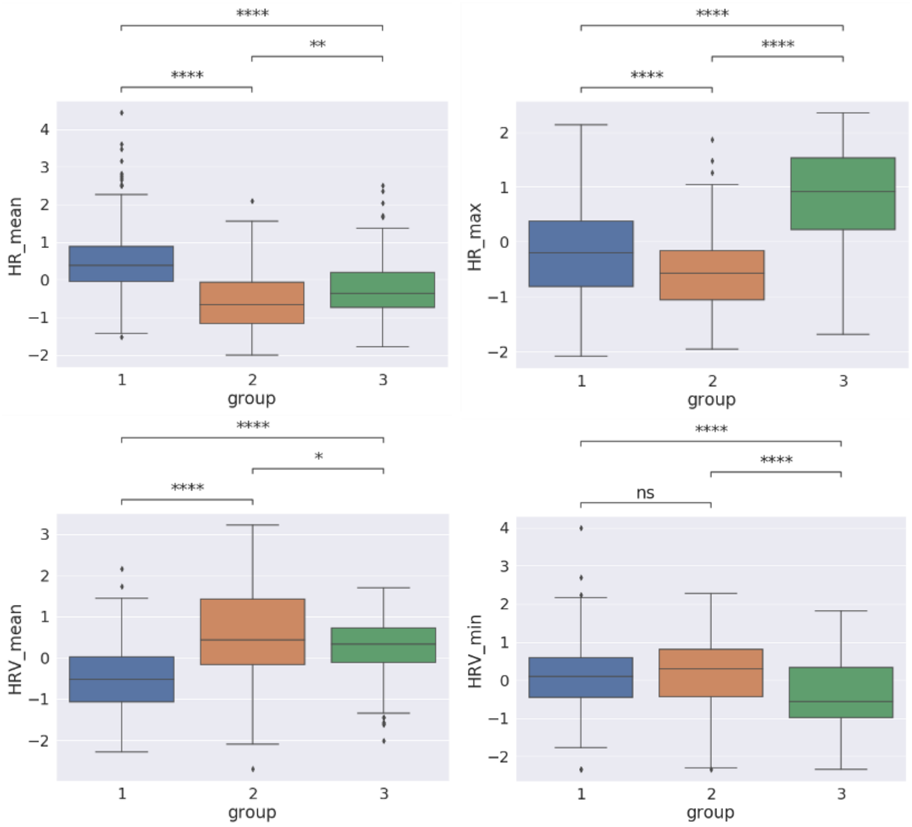

Figure 15: Selected boxplots and statistical significance differences of HR and HRV features for the three clustered groups.

Table 10: Mean ± standard deviation values of the daily averages of the 11 numerical sensor parameters for the three clustered groups. HR: heart rate; HRV: heart rate variability; RESP: respiratory rate; (m)s: (milli)seconds; bpm: beats per minute; C: Celsius degree; kOhm: kilo Ohm; mbar: millibar.

Feature	Unit	Group 1	Group 2	Group 3
ActivityCounts_mean	-	1.96 ± 0.36	1.81 ± 0.35	2.58 ± 0.63
Steps_mean	-	4.42 ± 2.62	4.13 ± 2.37	7.67 ± 4.42
EnergyExpenditure_mean	calories/s	1790.74 ± 484.72	1587.58 ± 372.15	2456.64 ± 1224.45
HR_mean	bpm	74.60 ± 6.54	67.13 ± 5.09	69.29 ± 5.62
HRV_mean	ms	40.53 ± 7.05	50.07 ± 10.07	47.01 ± 6.61
RESP_mean	bpm	17.37 ± 1.74	16.08 ± 1.10	15.41 ± 1.05
BloodPerfusion_mean	-	0.47 ± 0.18	0.57 ± 0.15	0.62 ± 0.25
BloodPulseWave_mean	-	2.62 ± 0.43	2.25 ± 0.28	2.15 ± 0.24
SkinTemperature_mean	C	34.15 ± 1.13	34.67 ± 0.92	33.68 ± 0.80
GalvanicSkinResponse_mean	kOhm	2.32 ± 0.58	1.97 ± 0.54	2.08 ± 0.55
Barometer_mean	mbar	979.18 ± 10.54	986.27 ± 9.40	983.94 ± 17.02

References

1 https://www.biovotion.com/everion/

2 https://www.biovotion.com/everion-precision-recognised-yet-again/

3 Barrios L, Oldrati P, Santini S, Lutterotti A. Recognizing digital biomarkers for fatigue assessment in patients with multiple sclerosis. In 12th EAI International Conference on Pervasive Computing Technologies for Healthcare–Demos, Posters, Doctoral Colloquium 2018. European Alliance for Innovation (EAI).

4 https://support.biofourmis.com/hc/en-us/categories/201377109-Everion-Device-

5 Tlija A, Węgrzyn-Wolska K, Istrate D. Missing-data imputation using wearable sensors in heart rate variability. Bulletin of the Polish Academy of Sciences: Technical Sciences. 2020:255-61.

6 Zhang J, Mu X, Fang J, Yang Y. Time Series Imputation via Integration of Revealed Information Based on the Residual Shortcut Connection. IEEE Access. 2019 Jul 15;7:102397-405.1

7 Cao W, Wang D, Li J, Zhou H, Li L, Li Y. Brits: Bidirectional recurrent imputation for time series. In Advances in Neural Information Processing Systems 2018 (pp. 6775-6785).

8 Franceschi JY, Dieuleveut A, Jaggi M. Unsupervised scalable representation learning for multivariate time series. In Advances in Neural Information Processing Systems 2019 (pp. 4650-4661).

9 https://scikit-learn.org/0.16/modules/generated/sklearn.cross_validation.StratifiedKFold.html

10 King G, Zeng L. Logistic Regression in Rare Events Data. Political Analysis. 2001;9(2):137–63.