

<https://doi.org/10.1038/s41746-024-01247-w>

Deep behavioural representation learning reveals risk profiles for malignant ventricular arrhythmias

Check for updates

Maarten Z. H. Kolk^{1,2}, Diana My Frodi³, Joss Langford^{4,5}, Tariq O. Andersen⁶, Peter Karl Jacobsen³, Niels Risum³, Hanno L. Tan^{1,7}, Jesper Hastrup Svendsen^{3,8}, Reinoud E. Knops^{1,2}, Søren Zöga Diederichsen³ & Fleur V. Y. Tjong^{1,2}✉

We aimed to identify and characterise behavioural profiles in patients at high risk of SCD, by using deep representation learning of day-to-day behavioural recordings. We present a pipeline that employed unsupervised clustering on low-dimensional representations of behavioural time-series data learned by a convolutional residual variational neural network (ResNet-VAE). Data from the prospective, observational SafeHeart study conducted at two large tertiary university centers in the Netherlands and Denmark were used. Patients received an implantable cardioverter-defibrillator (ICD) between May 2021 and September 2022 and wore wearable devices using accelerometer technology during 180 consecutive days. A total of 272 patients (mean age of 63.1 ± 10.2 years, 81% male) were eligible with a total sampling of 37,478 days of behavioural data (138 ± 47 days per patient). Deep representation learning identified five distinct behavioural profiles: Cluster A ($n = 46$) had very low physical activity levels and a disturbed sleep pattern. Cluster B ($n = 70$) had high activity levels, mainly at light-to-moderate intensity. Cluster C ($n = 63$) exhibited a high-intensity activity profile. Cluster D ($n = 51$) showed above-average sleep efficiency. Cluster E ($n = 42$) had frequent waking episodes and poor sleep. Annual risks of malignant ventricular arrhythmias ranged from 30.4% in Cluster A to 9.8% and 9.5% for Clusters D-E, respectively. Compared to low-risk profiles (D-E), Cluster A demonstrated a three-to-four fold increased risk of malignant ventricular arrhythmias adjusted for clinical covariates (adjusted HR 3.63, 95% CI 1.54–8.53, $p < 0.001$). These behavioural profiles may guide more personalised approaches to ventricular arrhythmia and SCD prevention.

Malignant ventricular arrhythmias are a main cause of sudden cardiac death (SCD)¹. In individuals at increased risk of SCD, patterns in physical behaviour (such as physical activity levels, sedentary behaviours, sleep behaviour) have emerged as potential prognostic indicators for ventricular arrhythmia onset, heart failure progression, and patient-reported outcomes^{2–5}. Wearable accelerometers provide a means for continuous measurement of these day-to-day physical behaviours in free-living environments⁶. Identifying patterns or clusters within behavioural time-

series data requires dimensionality reduction, as traditional clustering algorithms are unable to effectively process the granularity of such complex datasets. The process of dimensionality reduction, for instance reducing data to summary measures, may lead to the loss of intricate, non-linear associations in the data. Alternatively, deep neural networks are capable of learning low-dimensional latent representations from these complex datasets, while preserving the richness and intrinsic information present in the data^{7,8}. Unsupervised machine learning algorithms can then operate on

¹Department of Clinical and Experimental Cardiology, Amsterdam UMC Location University of Amsterdam, Heart Center, Meibergdreef 9, Amsterdam, the Netherlands. ²Amsterdam Cardiovascular Sciences, Heart Failure & Arrhythmias, Amsterdam UMC location AMC Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands. ³Department of Cardiology, Copenhagen University Hospital Rigshospitalet, Inge Lehmanns Vej 7, 2100 Copenhagen, Denmark. ⁴Activinsights Ltd., Unit 11, Harvard Industrial Estate, Kimbolton, Huntingdon, PE28 0NJ, United Kingdom. ⁵College of Life and Environmental Sciences, University of Exeter, Stocker Rd, Exeter, EX4 4PY, United Kingdom. ⁶Department of Computer Science, University of Copenhagen, Universitetsparken 1, 2100 Copenhagen, Denmark. ⁷Netherlands Heart Institute, Moreelsepark 1, 3511 EP Utrecht, The Netherlands. ⁸Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark. ✉e-mail: f.v.tjong@amsterdamumc.nl

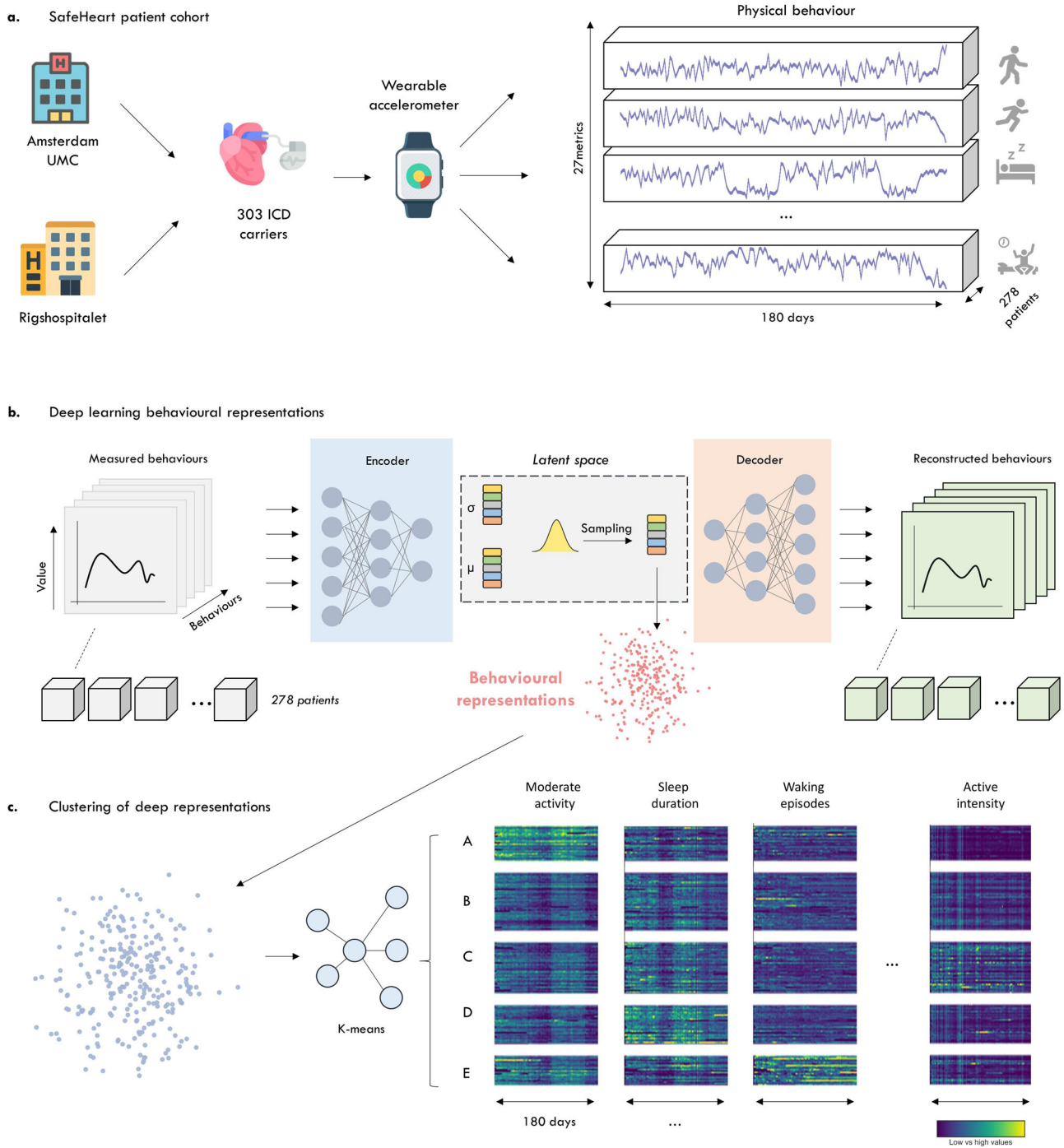


Fig. 1 | Workflow of the study. The workflow of the study is illustrated, that includes recruitment, data collection through a wearable device and data processing to identify behavioural profiles. **a** Recruitment of 303 patients from two international sites, who wore a wearable accelerometer for 180 consecutive days, during which behavioural metrics were recorded. **b** A convolutional residual variational

autoencoder learned the latent behavioural representations. The models was provided with time-series data for each subject that consisted of 27 variables measured at 180 timepoints. **c** Unsupervised clustering using a k-means algorithm of the behavioural representations identified distinct behavioural profiles.

these latent space representations to categorise similar samples into one cluster^{9,10}.

In this study, we aimed to identify and characterise behavioural profiles through deep representation learning in patients at risk of malignant ventricular arrhythmias (Fig. 1). Patients with an implantable cardioverter-defibrillator (ICD) were followed for six consecutive months using a wearable accelerometer to continuously monitor physical behaviour. A deep neural network was trained to learn a compressed representation from the behavioural time-series

data while preserving the relevant information. We hypothesised that through the clustering of these deep behavioural representations, we would be able to identify clinically meaningful behavioural profiles. These profiles were evaluated for their clinical relevance and association with the risk of ventricular arrhythmia.

Results

A total of 303 participants were enrolled in SafeHeart, of which 272 met the eligibility criteria for this study (21 patients did not wear the

Table 1 | Baseline characteristics in the total patient cohort and across behavioural profiles

	Cohort average (n = 272)	Cluster A (n = 46)	Cluster B (n = 70)	Cluster C (n = 63)	Cluster D (n = 51)	Cluster E (n = 42)	p
Age, years (SD)	63.1 (10.4)	62.8 (11.1)	61.9 (10.0)	63.7 (9.6)	64.7 (9.3)	62.7 (12.7)	0.640
Male, yes (%)	220 (80.9)	38 (82.6)	55 (78.6)	51 (81.0)	41 (80.4)	35 (83.3)	0.973
Secondary prevention ICD indication, n (%)	187 (68.8)	32 (69.6)	46 (65.7)	43 (68.3)	40 (78.4)	26 (61.9)	0.484
Body mass index, kg/m ² (SD)	28.2 (6.2)	29.2 (10.3)	27.4 (5.3)	27.9 (4.0)	28.8 (5.0)	28.0 (5.4)	0.598
CRT, n (%)	50 (18.4)	9 (19.6)	15 (21.4)	10 (15.9)	8 (15.7)	8 (19.0)	0.910
Heart disease, n (%)							
ICM	133 (48.9)	21 (45.7)	29 (41.4)	31 (49.2)	28 (54.9)	24 (57.1)	0.457
DCM	45 (16.5)	7 (15.2)	13 (18.6)	14 (22.2)	5 (9.8)	6 (14.3)	0.467
HCM	11 (4.0)	0 (0)	5 (7.1)	3 (4.8)	1 (2.0)	2 (4.8)	0.357
Cardiovascular history, n (%)							
Myocardial infarction	101 (37.1)	20 (43.5)	21 (30.0)	24 (38.1)	21 (41.2)	15 (35.7)	0.603
Heart failure (HFrEF)	147 (54.0)	24 (52.2)	39 (55.7)	33 (52.4)	26 (51.0)	25 (59.5)	0.923
Diabetes mellitus	45 (16.5)	5 (10.9)	13 (18.6)	11 (17.5)	10 (19.6)	6 (14.3)	0.769
Hypertension	136 (50.0)	19 (41.3)	36 (51.4)	27 (42.9)	35 (68.6)	19 (45.2)	0.037
Atrial fibrillation	95 (34.9)	13 (28.3)	23 (32.9)	22 (34.9)	23 (45.1)	14 (33.3)	0.493
Medication, n (%)							
ACE inhibitor	110 (40.4)	14 (30.4)	22 (31.4)	31 (49.2)	25 (49.0)	18 (42.9)	0.094
Angiotensin receptor blocker	66 (24.3)	10 (21.7)	19 (27.1)	14 (22.2)	13 (25.5)	10 (23.8)	0.956
Loop diuretics	90 (33.1)	17 (37.0)	20 (28.6)	20 (31.7)	15 (29.4)	18 (42.9)	0.536
β-blocker	219 (80.5)	33 (71.7)	57 (81.4)	51 (81.0)	43 (84.3)	35 (83.3)	0.560
Lipid lowering drugs	174 (64.0)	26 (56.5)	49 (70.0)	39 (61.9)	34 (66.7)	26 (61.9)	0.633

ACE Angiotensin-Converting Enzyme, CRT Cardiac Resynchronisation Therapy, ICM Ischaemic cardiomyopathy, DCM Dilated cardiomyopathy, HCM Hypertrophic cardiomyopathy, HFrEF heart failure with reduced ejection fraction.

GENEActiv wearable accelerometer, 10 patients did not meet the required minimum of 30 days of behavioural data). A total of 37,478 days of wearable data were collected (mean 138 ± 47 days per patient). Table 1 shows the clinical characteristics of the patient cohort. Patients had a mean age of 63.1 ± 10.4 years and 80.9% were male, 133 (48.9%) patients had ischaemic heart disease as cause of heart failure, 147 (54.0%) had heart failure with reduced ejection fraction (HFrEF), and 187 (68.8%) had a secondary prevention ICD indication. Fifty (18.4%) patients received cardiac resynchronisation therapy (CRT), the majority of patients used a β-blocker (80.5%). All patients completed one-year of follow-up, during which 46 (16.9%) patients received appropriate ICD therapy for a malignant ventricular arrhythmia, five (1.8%) patients received inappropriate ICD therapy and four (1.4%) patients died.

Characteristics of the identified behavioural profiles

Five behavioural profiles were identified: A (n = 46), B (n = 70), C (n = 63), D (n = 51) and E (n = 42). Mean values for behavioural metrics across the clusters are displayed in Table 2, Supplementary Fig. 4 provides a granular representation of behaviours during the 180-day monitoring period. Figure 2 shows the individual behavioural metrics for each profile, relative to the cohort averages. In summary, Clusters B and C were characterised by active profiles with high daily steps counts (15353 ± 4062 & 13577 ± 4130 steps). The volume of activity (2921 ± 1410 gs) in Cluster B was accumulated over a longer period (370 ± 76 min) and at a lower average intensity (123 ± 17 mg), while the volume of activity (2527 ± 810 gs) in Cluster C was linked to a greater number of faster walking steps (3689 ± 2100 steps) at higher cadences (92 ± 10 steps/min) and intensities (132 ± 22 gs). Clusters D and E had less active profiles with fewer daily steps (9971 ± 3103 & 9291 ± 3378 steps). Cluster D had the longest inactive bout durations and fewest

number per day (0.72 ± 0.25 mins and 632 ± 110). The behavioural patterns of Cluster E were more fragmented with shorter active bouts (0.43 ± 0.12 min) and more inactive bouts (863 ± 135). Cluster A had the highest inactive duration (835 ± 110 min), the lowest activity intensity (103 ± 33 mg) and least number of steps (6246 ± 2406 steps). In the sleep domain, Cluster A was characterised by the shortest total sleep duration (281 ± 86 min), the longest average duration of wake after sleep onset interruptions (6.8 ± 2.1 min) and the longest sleep onset latency (10.8 ± 5.1 min). The nocturnal patterns of Cluster E were fragmented, similar to the day, with the lowest sleep efficiency (51.5 ± 8.6%), most wake after sleep onset interruptions (32 ± 8) and shortest maximum sleep bout lengths (42 ± 10 min). Cluster D had the long sleep interval and total sleep durations (611 ± 122 and 369 ± 75 min) with longest maximum sleep bout lengths (57 ± 11 min). The sleep profiles of Clusters B and C were unexceptional other than Cluster B having the shortest sleep interval duration (507 ± 89 min).

Cluster characterisation

SHAP values were computed to represent the most important behavioural markers that characterised each behavioural profile. Figure 3a shows the variables with highest feature importance across behavioural profiles. The duration spent in moderate activity, the amount of slow steps and the number of sleep events were the behavioural markers that differentiated most between clusters. Figure 3b–f illustrates the top behavioural features that predict membership of each of the clusters. Clusters C and E were predicted by a combination of sleep behaviours and movement behaviours, while the other clusters were predominantly predicted by movement behaviours alone. No statistically significant differences were observed between clusters in terms of medication usage and medical history, apart from hypertension (p = 0.037) (Table 1).

Table 2 | Mean values of each behavioural metric collected over a 180-day period in the total patient cohort and across behavioural profiles

	Cohort average (n = 272)	Cluster A (n = 46)	Cluster B (n = 70)	Cluster C (n = 63)	Cluster D (n = 51)	Cluster E (n = 42)	p
Movement behaviour, mean (SD)							
Mean cadence (step/min)	54.3 (4.7)	53.0 (4.8)	54.0 (3.6)	57.9 (4.8)	52.6 (4.7)	52.8 (3.2)	<0.001
Cadence 95p (step/min)	80.4 (12.3)	75.7 (13.0)	78.6 (10.0)	91.6 (9.8)	77.4 (10.5)	75.7 (10.7)	<0.001
Fast steps (count)	1976.6 (1818.9)	847.5 (1115.8)	2141.6 (1650.5)	3689.2 (2099.6)	1313.4 (886.0)	1174.6 (1049.8)	<0.001
Slow steps (count)	5233.2 (3209.5)	1805.6 (884.1)	8061.9 (3088.4)	5749.1 (2757.2)	4477.7 (2152.9)	4416.1 (2320.6)	<0.001
Total steps (count)	11456.4 (4825.3)	6246.3 (2406.4)	15352.8 (4061.7)	13577.3 (4130.1)	9971.3 (3102.5)	9290.8 (3377.5)	<0.001
Activity volume (gs)	2215.2 (1596.7)	1550.3 (2984.1)	2920.5 (1410.4)	2527.4 (810.4)	1836.4 (586.7)	1759.9 (655.9)	<0.001
Vigorous activity (min)	3.0 (18.1)	5.3 (35.0)	3.1 (20.5)	4.2 (7.1)	0.78 (2.2)	1.05 (1.7)	0.694
Moderate activity (min)	129.5 (74.1)	46.1 (23.6)	187.6 (69.9)	154.3 (61.5)	113.2 (52.1)	106.7 (54.9)	<0.001
Light activity (min)	150.9 (43.7)	134.9 (45.8)	179.1 (42.0)	148.2 (37.6)	143.9 (40.4)	134.1 (35.1)	<0.001
Active duration (min)	283.4 (94.4)	186.3 (62.0)	369.8 (76.2)	306.6 (72.3)	257.9 (67.7)	241.8 (68.4)	<0.001
Inactive duration (min)	718.2 (113.5)	835.4 (110.4)	670.6 (90.7)	692.5 (96.6)	667.2 (89.1)	769.4 (88.8)	<0.001
M6 intensity (g)	228.9 (78.8)	165.6 (54.5)	243.2 (53.8)	290.5 (96.2)	203.0 (49.0)	213.5 (65.1)	<0.001
Active intensity (g)	118.6 (22.9)	103.4 (32.6)	122.6 (17.3)	131.5 (21.6)	112.9 (13.7)	115.9 (15.9)	<0.001
Active event duration (min)	0.62 (0.21)	0.60 (0.29)	0.74 (0.18)	0.64 (0.16)	0.64 (0.15)	0.43 (0.12)	<0.001
Active event (count)	215.6 (53.56)	159.8 (45.5)	245.3 (38.8)	229.4 (41.3)	194.6 (51.3)	231.7 (47.7)	<0.001
Inactive event duration (min)	0.64 (0.13)	0.6 (0.05)	0.61 (0.06)	0.62 (0.05)	0.72 (0.25)	0.61 (0.06)	<0.001
Inactive events (count)	758.5 (153.5)	887.3 (148.7)	726.1 (122.6)	733.2 (116.6)	631.6 (109.6)	863.4 (134.8)	<0.001
Sleep behaviour, mean (SD)							
Sleep onset latency (min)	7.6 (4.7)	10.8 (5.1)	6.7 (3.2)	6.7 (3.5)	8.9 (6.9)	5.6 (2.3)	<0.001
Sleep interval duration (min)	549.9 (107.7)	541.4 (105.4)	506.9 (89.1)	531.3 (83.5)	610.7 (122.2)	585.5 (112.9)	<0.001
Total sleep duration (min)	313.0 (77.3)	280.8 (86.2)	294.2 (62.1)	322.9 (62.0)	368.5 (74.6)	297.2 (78.9)	<0.001
Sleep efficiency (%)	56.7 (8.9)	52.5 (11.2)	57.7 (7.1)	59.9 (6.8)	59.5 (8.2)	51.5 (8.6)	<0.001
WASO (count)	23.7 (7.2)	21.0 (7.1)	21.9 (5.2)	21.8 (4.9)	24.0 (6.0)	32.1 (8.2)	<0.001
WASO duration (min)	5.7 (1.9)	6.8 (2.1)	5.7 (1.6)	5.2 (1.3)	5.6 (2.4)	5.5 (1.9)	<0.001
Time first WASO (min)	9.6 (5.4)	9.9 (3.5)	10.3 (7.2)	9.5 (4.0)	10.1 (7.0)	7.3 (2.3)	0.052
Longest sleep period (min)	50.2 (11.8)	49.9 (12.9)	49.4 (11.3)	51.7 (10.4)	56.6 (10.8)	42.1 (9.7)	<0.001
Naps duration (min)	11.9 (11.5)	7.7 (4.6)	9.5 (4.7)	13.1 (5.9)	16.4 (21.5)	13.1 (12.0)	0.001
Sleep events (count)	107.3 (25.1)	106.7 (18.7)	98.3 (21.1)	97.7 (18.0)	108.1 (20.3)	136.2 (29.9)	<0.001

min minutes, M6 most active 6 minutes, SD standard deviation, WASO wake up after sleep onset.

Patient-reported outcomes across behavioural profiles

A total of 239 patients filled out questionnaires at the study baseline (non-response rates for subsequent clusters A-E were 11.4%, 15.6%, 12.7%, 9.5%, and 15.2%, respectively). Median scores for the EQ-5D-5L and KCCQ domains are provided in Supplementary Table 2. In particular, Cluster A reported physical limitations, Cluster C reported high self-efficacy but worse social limitations, Cluster D highest disease-specific quality of life, and Cluster E reported highest burden of symptoms (Supplementary Fig. 5). Differences in patient-reported outcomes between clusters were not statistically significant.

Incidence of the outcomes of interest across behavioural profile

Figure 4a shows the risk of malignant ventricular arrhythmias treated by the ICD across the clusters during one-year follow-up. Event rates for clusters A until E were respectively 30.4%, 17.1%, 17.5%, 9.8% and 9.5% (log-rank *p* value 0.06). As displayed in Fig. 4b, the risk of malignant ventricular arrhythmias was significantly higher in Cluster A (unadjusted HR 2.26, 95% CI 1.20–4.23, *p* = 0.01), which remained after adjusting for clinical covariates (adjusted HR 2.30, 95% CI 1.21–4.36, *p* = 0.01). Also, the risk of

malignant ventricular arrhythmias in the low-risk behavioural profiles (Cluster D-E) was significantly lower compared to the other clusters (unadjusted HR 0.45, 95% CI 0.22–0.94, *p* = 0.03). Inappropriate ICD therapy was delivered in three patients in Cluster A (4.3%), two patients in Cluster B (2.9%), and one patient in Cluster C (1.6%). In total four patients died during follow-up, of which two in Cluster D (3.9%), one in Cluster A (2.2%) and one in Cluster C (1.6%). A significant difference in the composite endpoint between clusters was observed (log-rank *p* value 0.04) (Fig. 4c). Unadjusted and adjusted hazard ratios for the respective clusters for the composite endpoint are displayed in Table 3. In Supplementary Fig. 6, ROC curves of logistic regression models predicting cases of malignant ventricular arrhythmias and the composite endpoint are presented. Regression models that included cluster membership within their feature set demonstrated superior performance, compared to models that excluded this variable.

Discussion

In this study, we demonstrated deep representation learning of complex day-to-day movement and sleep behaviours to enable the identification



Fig. 2 | Average values for physical behaviour measurements across the behavioural profiles. The average values for the behavioural measurements for each behavioural profile are displayed. **a** The bar charts displayed on the left depict the average values for the metrics that reflect movement behaviour across behavioural

profiles, relative to the cohort average. **b** Bar charts depicted on the right display the average values for the metrics that reflect sleep behaviours across the behavioural profiles. All values were scaled using z-scores.

of clinically relevant behavioural profiles. These profiles were associated with an annual risk of malignant ventricular arrhythmias ranging from 30.4% to 9.5%. Our research extends prior work, bringing forth two novelties. First, while prior studies have evaluated physical behavioural metrics over monitoring intervals up to 14 days³, we identified distinct behavioural profiles derived from continuous accelerometer

measurements spanning six months. Second, earlier studies have mainly focused on individual metrics for activity or sleep, despite these 24-hour rest-activity behaviours being highly interrelated. In the present work, we took a more holistic approach to physical behaviour by modelling the interplay between various concurrent behavioural mechanisms and their potential implications for clinical events.

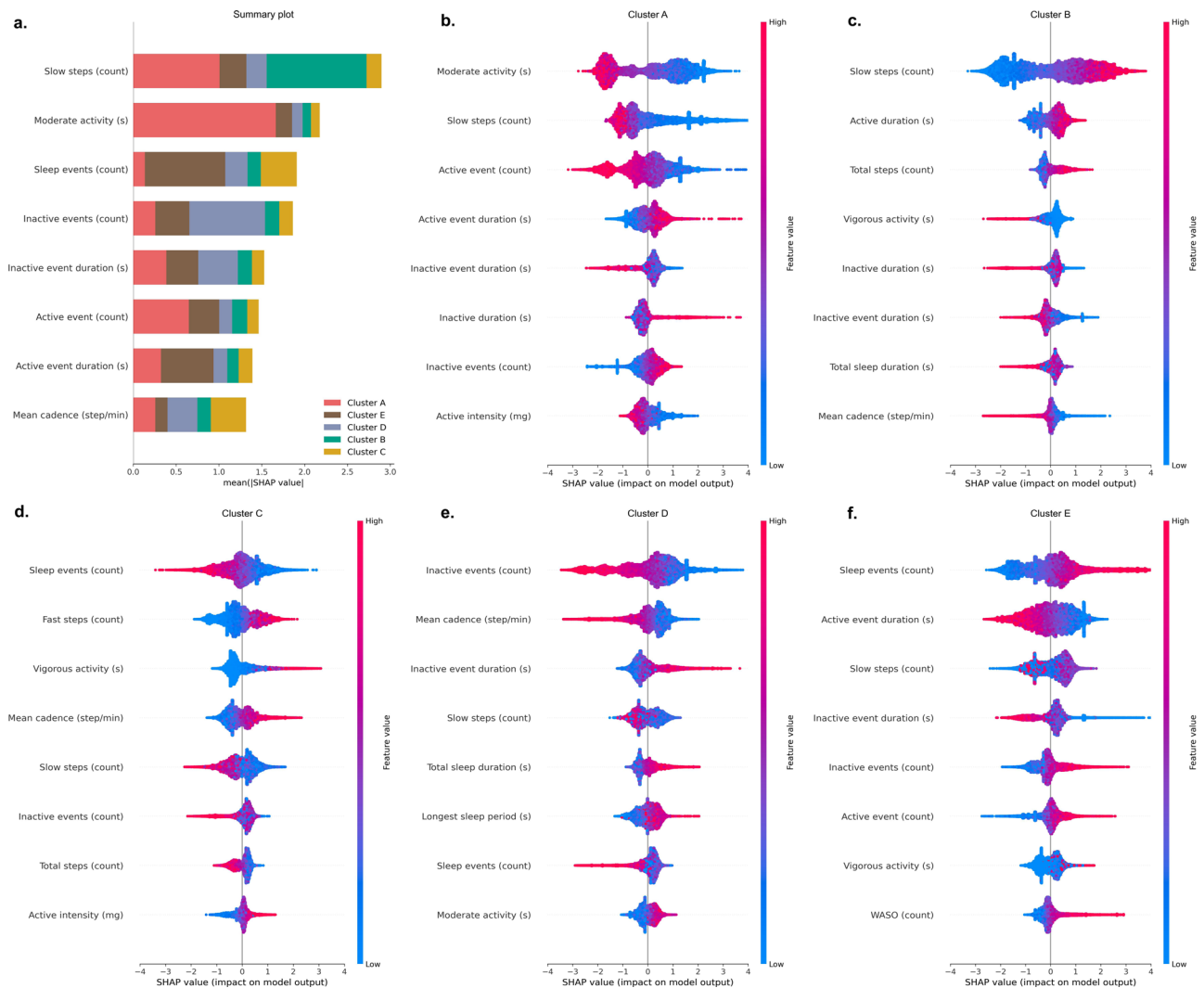


Fig. 3 | Characterisation of the behavioural profiles through Shapley values obtained from a trained machine learning model. A trained machine learning classifier (extreme gradient boosting) was used to predict membership of a profile based on daily behavioural measurements. **a** The bar chart represents the importance of features used by the extreme gradient boosting model to predict each profile. Horizontal bars represent the average contribution of a behavioural metric for the

predicted profile. The features are ranked based on the summed importance of that feature to predict each profile. **b–f** The SHAP summary plot is displayed for each behavioural profile. The features are ranked by the mean absolute SHAP value. A positive SHAP value suggests a positive contribution, while a negative value indicates a negative contribution. The model predicted membership of the behavioural profile based on the daily measurements with an AUROC of 0.99.

Despite considerable variations in clinical trajectories among patients with an ICD, current follow-up strategies remain *one-size-fits-all*. Advances in wearable technologies have removed barriers for the continuous measurement of behavioural patterns, which could make wearables suitable as screening tools to identify individuals at-risk of disease progression. Several studies have shown that continuous activity measurements could indicate a decline in functional status, progression of heart failure, or onset of atrial fibrillation, each potentially increasing the risk of ventricular arrhythmia onset^{3,4,11,12}. However, physical activity is also a modifiable risk factors that may reduce ventricular arrhythmia risk by alteration of autonomic tone, mitigation of the catecholamine release observed during exercise and an increase of resting parasympathetic tone¹³. Recent analyses of data from the UK Biobank have demonstrated a reduction in the risk of ventricular arrhythmia amongst physically active individuals^{14,15}. With data from this prospective study, we demonstrated that an active behavioural profile does not necessarily reduce the risk of ventricular arrhythmia, which highlights the importance of considering various behaviours simultaneously. In particular, Clusters B and C had annual event rates of ~17% despite their daily time spent

physically active being substantially higher compared to the other profiles. In contrast, Clusters D and E were half as likely to experience the outcome of interest, despite having less active profiles. This indeed suggests that interplays between various behaviours, such as intermittent sedentary behaviour with isolated bouts of physical exertion, rather than isolated measurements of activity characteristics, may explain differences in risk of ventricular arrhythmia onset. Furthermore, the absence of significant associations between patient-reported outcomes (e.g. symptom severity and physical limitations) and behavioural profiles might indicate that these are phenotypic in their origin rather than representing more transient behavioural patterns.

Our findings support the notion that abnormalities in 24-hour rest-activity patterns modulate the risk of ventricular arrhythmia onset. Circadian rhythm disruption has been associated with increased risk of atrial fibrillation onset¹⁶ and heart failure¹⁷ in previous studies. We observed an annual risk of ventricular arrhythmia exceeding 30% in the behavioural profile characterised by sedentary behaviour, a lack of high-intensity activity, and disturbed sleep behaviour. Adjusted for clinical covariates, this profile was associated with a three-to-four fold risk of experiencing a

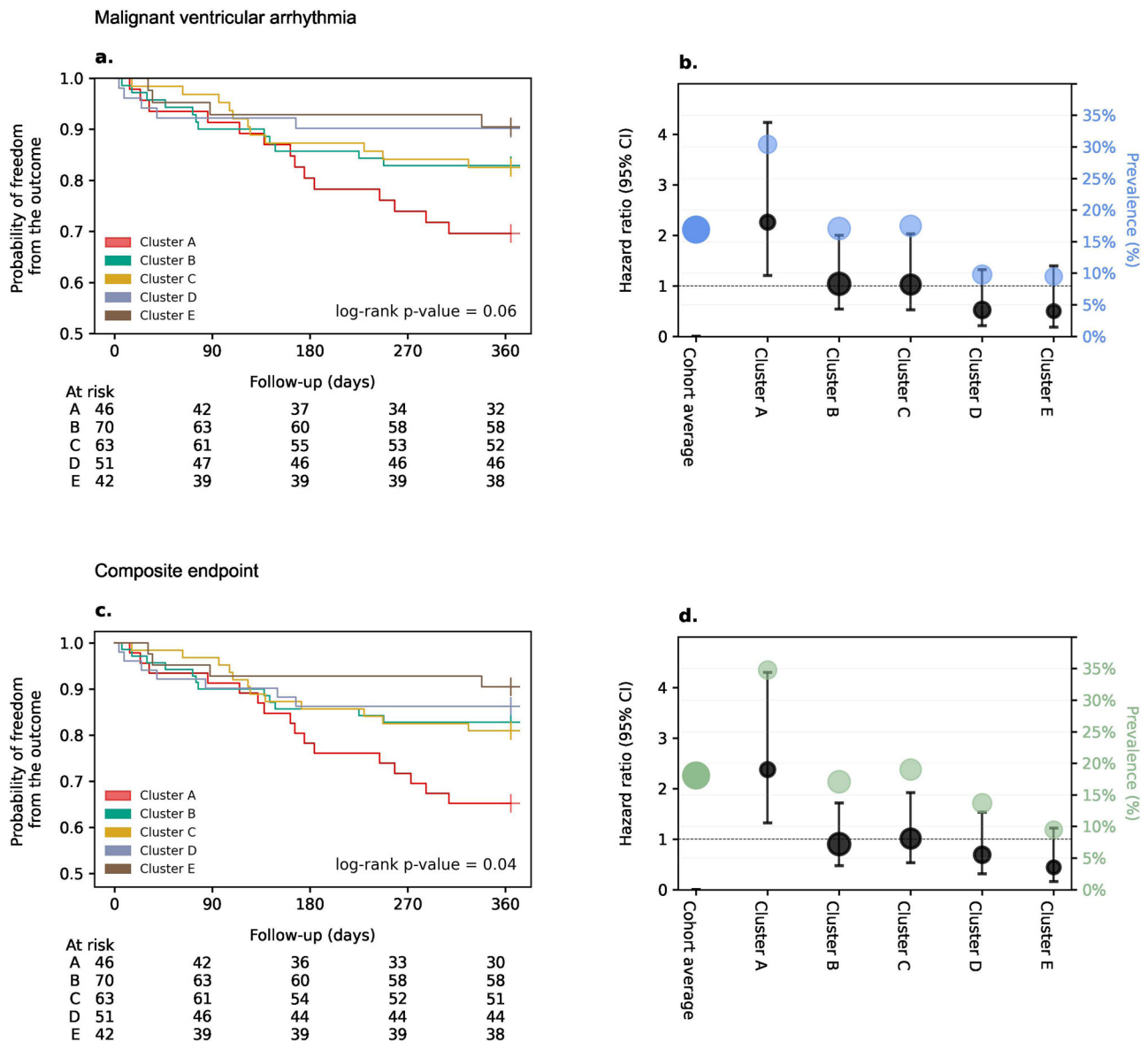


Fig. 4 | Comparative results of the outcomes of interest for patients stratified by the behavioural profiles. Time-to-event analyses according to the behavioural profiles are presented. **a** Kaplan-Meier curves for malignant ventricular arrhythmias treated by the ICD, and **b** hazard ratios and 95% confidence intervals obtained from the Cox proportional-hazards model. **c** Kaplan-Meier curves for the composite endpoint of all ICD therapy and mortality, and **d** hazard ratios and 95% confidence intervals. The prevalence of the outcome is displayed as a percentage, represented by the blue and green circles. Distributions of times to events were compared with the log-rank test.

Table 3 | Associations of the five behavioural profiles with malignant ventricular arrhythmias treated by the ICD, and the composite of all ICD therapies and mortality

	Malignant ventricular arrhythmia				Composite endpoint			
	Unadjusted HR (95% CI)	p	Adjusted HR (95% CI) ^a	p	Unadjusted HR (95% CI)	p	Adjusted HR (95% CI) ^a	p
Cluster A	2.26 (1.20–4.23)	0.01	2.30 (1.21–4.36)	0.01	2.38 (1.32–4.30)	<0.001	2.46 (1.35–4.49)	<0.001
Cluster B	1.04 (0.54–2.00)	0.91	1.13 (0.58–2.20)	0.72	0.90 (0.47–1.71)	0.44	0.99 (0.51–1.90)	0.97
Cluster C	1.03 (0.52–2.03)	0.93	1.03 (0.52–2.02)	0.94	1.01 (0.53–1.92)	0.98	1.00 (0.53–1.92)	0.99
Cluster D	0.52 (0.21–1.32)	0.17	0.53 (0.20–1.35)	0.18	0.69 (0.31–1.53)	0.36	0.66 (0.29–1.49)	0.32
Cluster E	0.50 (0.18–1.39)	0.18	0.43 (0.15–1.22)	0.11	0.44 (0.16–1.22)	0.12	0.39 (0.14–1.09)	0.07

CI confidence interval, HR hazard ratio.

^aAdjusted for age, sex, indication for ICD implantation, and presence of atrial fibrillation, heart failure and cardiac resynchronisation therapy.

ventricular arrhythmia compared to the low-risk profiles. While these findings should be validated in larger cohorts, they emphasise the importance of comprehensive modelling of physical behaviour. The use of wearable devices for behavioural profiling holds promise for follow-up strategies tailored to an individual patient.

Clustering of deep learning-derived latent representations comes with the limitation of interpretability, as the latent representations are inferred from the underlying data and are not directly explainable (*black box*). To provide transparency, we characterised clusters by assessment of feature importance of a trained machine learning classifier that predicts cluster memberships based on day-to-day behavioural metrics¹⁸. A second limitation to our study is the use of processed output from the accelerometer, instead of the underlying raw accelerometry output. Some of these metrics are created through application of specific thresholds that rely on the calibration studies, but pose a challenge when comparing metrics among different studies or populations¹⁹. Third, from our findings, it remains uncertain whether the behavioural profiles can be generalised to other populations, such as heart failure patients who do not satisfy the criteria for an ICD, and thus warrant future research. Fourth, despite cluster membership showing significant associations with the outcome of interest after adjusting for clinical covariates, there is a risk of residual confounding. For instance, high-risk behavioural profile (Cluster A) was characterised by higher proportion of patients with a prior myocardial infarctions; however, these patients did not receive prescriptions for β -blockers, lipid-lowering drugs, or ACE inhibitors. This could point towards a potential undertreatment of these patients. This study was entirely decentralised in its design, with patient recruitment, informed consent, and study procedures conducted without physical contact between the study staff and participants. Consequently, information from imaging modalities (e.g., LVEF) and electrocardiography at the time of enrolment was not available. Future studies exploring the interplay between these clinical patient characteristics, and behavioural profiles are warranted.

Deep representation learning of physical behavioural patterns identifies distinct behavioural profiles with significant differences in their risk of malignant ventricular arrhythmia and death. Behavioural profiling using objective and real-time measurements obtained from wearable devices may enable clinicians to adjust and optimise treatment and prevention strategies to an individual patient. Interpretability of clustered latent representations and relatively small sample sizes prompt the need for further investigation into the mechanisms underlying their influence on ventricular arrhythmia risk and SCD.

Methods

Ethics

The study was approved by the Institutional Review Boards of the Amsterdam University Medical Center (date 09-04-2021, approval number 2020/248) and Copenhagen University Hospital Rigshospitalet (date 19-04-2021, approval number H-20081068). All participants provided informed consent prior to their enrolment. The study was conducted in accordance with the Declaration of Helsinki.

Study design and setting

This is an analysis of the international SafeHeart study, a prospective, observational study conducted at two tertiary academic centers in Europe (Amsterdam University Medical Center, the Netherlands and Copenhagen University Hospital Rigshospitalet, Denmark). The purpose of this study was to develop a personalised model to predict ICD therapy for malignant ventricular arrhythmia²⁰. Data used to create the prediction model included recordings from a wearable accelerometry recording device. Patient inclusion was conducted through telephone-based procedures between May 2021 and September 2022, the enrolment date was defined as the day when the wearable device was delivered to the patient. Throughout the study, participants had the option to withdraw from the study at any stage, either partially (by discontinuing the use of the wearable device) or completely. We adhered to the *Strengthening the Reporting of Observational Studies in*

Epidemiology (STROBE) reporting guidelines for observational studies²¹. The study was registered at the National Trial Registration in the Netherlands (Trial NL9218; <https://www.onderzoekmetmensen.nl/en>).

Participants

Participants qualified for enrolment in the SafeHeart study if they fulfilled the following conditions: i) received an ICD with or without cardiac resynchronisation therapy (CRT-D) in the five years leading up to enrolment, ii) experienced appropriate or inappropriate ICD therapy (high voltage shock therapy or anti-tachycardia pacing (ATP)) or demonstrated evidence of ventricular arrhythmias within eight years prior to enrolment, iii) engaged in a remote ICD monitoring programme, and iv) were at least 18 years old. Exclusion criteria were severe physical disability, end-stage heart failure, and a life expectancy of less than one year. The study protocol with the entire list of inclusion and exclusion criteria has been published previously²⁰.

Physical behaviour measurements

Accelerometer-based wearable devices allow for continuous and objective quantification of daily physical behaviour by the recording of body movement along reference axes and signal analysis (e.g., intensity, frequency, and volume of activity and postural changes). In this study, various behavioural metrics were collected including daily activity and inactivity durations, duration of activity and inactivity episodes, activity intensity, activity volume, step count (total, slow and fast), cadence, sleep duration, sleep efficiency, wake up after sleep onset (WASO), nap duration and sleep onset latency. A complete overview of the collected metrics and their definitions are displayed in Supplementary Table 1. To collect these metrics, participants wore the GENEActiv Original 1.1 accelerometer (Activinsights Ltd, Cambridgeshire, United Kingdom) on the wrist for 6 months. Devices were returned (for data extraction) and replaced biweekly or every 4 weeks. Continuous raw data were recorded at 50 Hz or 20 Hz and converted into daily summaries^{22,23}. Patients were eligible for this study if they had at least 30 days of wearable data.

Outcome of interest

The prospective collection of the outcomes of interest occurred at both sites from enrolment in the study onwards. These outcomes were: i) any malignant ventricular arrhythmia defined as an episode of sustained ventricular tachycardia or ventricular fibrillation, treated by the ICD through a shock and/or ATP; ii) a composite endpoint comprising all ICD therapies and death. ICD therapies encompassed those for malignant ventricular arrhythmias, in addition to those in response to rhythms other than sustained ventricular tachycardia or ventricular fibrillation (e.g. atrial fibrillation, sinus tachycardia).

Patient reported outcome measures

Two patient reported outcome measures (PROMs), the EuroQoL 5-Dimensions 5-Levels (EQ-5D-5L) and Kansas City Cardiomyopathy Questionnaire (KCCQ), were used in the SafeHeart study^{24,25}. Both PROMs were filled out by participants at study enrolment. The EQ-5D-5L assesses health across five domains, yielding a utility score ranging from -0.590 to 1.000 . Meanwhile, the KCCQ, designed for heart failure patients, provided scores on a scale of 0 to 100, subdivided into the domains symptom burden, physical limitation, social limitation, quality of life and self-efficacy.

Deep representation learning of physical behaviour data

We derived deep representations from the day-to-day behavioural time-series collected during the first six months of the study (Fig. 1a). Specifically, we used a β -variational autoencoder (VAE) that encodes input data through a probabilistic approach (mapping data into a probability distribution) and decodes from this distribution back into reconstructed data (Fig. 1b)^{26,27}. Supplementary Fig. 1 presents a schematic overview of the VAE architecture. The inputs were longitudinal trajectories of 27 behavioural metrics over 180 days, resulting

in an input dimension of $272 \times 27 \times 180$. Missing values of behavioural metrics were linearly interpolated and normalised. Our trained VAE reconstructed the behavioural time-series with a Pearson Correlation Coefficient of 0.988 ± 0.0379 , a root mean square error (RMSE) of 0.031 ± 0.026 and a percentage root-mean-square difference (PRD) of 10.553 ± 0.038 . Supplementary Fig. 2 depicts an example of the trends in behavioural measurements along with the reconstructed trend derived from 32 latent variables. The VAE models were developed using PyTorch (version 2.0.5) in Python (version 3.6.7).

We then applied an unsupervised machine learning algorithm to cluster these representations (Fig. 1c). The k-means algorithm aims to minimise the within-cluster variance, making data points within the same cluster as similar as possible and data points in different clusters as dissimilar as possible. The appropriate number of clusters was assessed using within-cluster variation (inertia), silhouette scores, and Davies-Bouldin index (Supplementary Fig. 3). Considering that the k-means algorithm operates stochastically, and initialisation of the model may affect the decision of the optimal k , we averaged the results over multiple iterations to reduce the impact of randomness²⁸. We evaluated cluster stability by computing the Jaccard index across 100 bootstrapped samples²⁹. Clustering was performed using the scikit-learn library (version 1.3.0)³⁰.

Cluster characterisation through cluster membership prediction

We aimed to characterise the identified clusters using SHapley Additive exPlanations (SHAP) values³¹. SHAP values are widely used to determine the contribution of particular features to the predicted outcome. We derived SHAP values from a supervised machine learning classifier (eXtreme Gradient Boosting), which was trained to predict cluster membership from daily behavioural values (48,960 days)³². Subsequent ranking of these SHAP values provides insight into behavioural metrics that contribute positively (or negatively) to cluster membership. We assessed the performance of these classifications using the receiver operating characteristic curve (ROC).

Statistical analysis

Continuous variables were presented by the median, mean, interquartile range, and standard deviation. Categorical socio-demographic and clinical variables were presented as frequencies (percentages) and compared using the χ^2 test. T-tests were used for pairwise comparisons, analysis of variance (ANOVA) for assessing differences among multiple groups with normally distributed data. The Mann-Whitney U test was used for non-normally distributed variables, and the Kruskal-Wallis test for comparisons involving more than two groups with non-normally distributed data. The risk of the outcomes of interest during follow-up was estimated using the Kaplan-Meier method; log-rank tests were used to compare survival between clusters. Cox Proportional Hazard models were used to assess the association between behavioural profiles and the risk of outcomes of interest. The model included the clinical covariates age, sex, indication for ICD implantation, presence of atrial fibrillation, heart failure, and type of ICD. Schoenfeld residuals were used to check the proportional hazards assumption. A two-sided p value < 0.05 was considered significant. The prognostic significance of the behavioural profiles for the outcomes of interest was assessed through logistic regression models. Two models were constructed for each outcome of interest: the first model included clinical patient information (medical history and medication status) along with cluster membership as input features, while the second model excluded cluster membership. Prediction accuracy was assessed through stratified k-fold cross-validation, and quantified using the area under the receiver operating characteristic curve (AUROC).

Data availability

Data sharing requests will be considered upon a reasonable request. For access, please email the corresponding author.

Code availability

Code scripts are available at: <https://github.com/DeepRiskAUMC/Deep-representation-clustering>.

Received: 12 January 2024; Accepted: 30 August 2024;

Published online: 16 September 2024

References

- Empana, J. P. et al. Incidence of sudden cardiac death in the European Union. *J. Am. Coll. Cardiol.* **79**, 1818–1827 (2022).
- Shim, J., Fleisch, E. & Barata, F. Wearable-based accelerometer activity profile as digital biomarker of inflammation, biological age, and mortality using hierarchical clustering analysis in NHANES 2011–2014. *Sci. Rep.* **13**, 9326 (2023).
- Kolk, M. Z. H. et al. Accelerometer-assessed physical behavior and the association with clinical outcomes in implantable cardioverter-defibrillator recipients: A systematic review. *Cardiovasc Digit Health J.* **3**, 46–55 (2022).
- Rosman, L., Lampert, R., Sears, S. F. & Burg, M. M. Measuring physical activity with implanted cardiac devices: a systematic review. *J Am Heart Assoc* **7**, e008663 (2018).
- Kolk, M. et al. Behavioural digital biomarkers enable real-time monitoring of patient-reported outcomes: a substudy of the multicenter, prospective observational SafeHeart study. *Eur Heart J Qual Care Clin Outcomes* qcad069, (2023).
- Hughes, A., Shandhi, M. M. H., Master, H., Dunn, J. & Brittain, E. Wearable devices in cardiovascular medicine. *Circ. Res* **132**, 652–670 (2023).
- Kolk, M. Z. H. et al. Machine learning of electrophysiological signals for the prediction of ventricular arrhythmias: systematic review and examination of heterogeneity between studies. *EBioMedicine* **89**, 104462 (2023).
- Kolk, M. Z. H. et al. Dynamic prediction of malignant ventricular arrhythmias using neural networks in patients with an implantable cardioverter-defibrillator. *EBioMedicine* **99**, 104937 (2023).
- Min, E. et al. A survey of clustering with deep learning: from the perspective of network architecture. *IEEE Access* **6**, 39501–39514 (2018).
- Karim, M. R. et al. Deep learning-based clustering approaches for bioinformatics. *Brief. Bioinform* **22**, 393–415 (2021).
- Ginder, C. et al. Predicting malignant ventricular arrhythmias using real-time remote monitoring. *J. Am. Coll. Cardiol.* **81**, 949–961 (2023).
- Shakibfar, S. et al. Predicting electrical storms by remote monitoring of implantable cardioverter-defibrillator patients using machine learning. *Europace* **21**, 268–274 (2019).
- Soares-Miranda, L. et al. Physical activity and heart rate variability in older adults: the Cardiovascular Health Study. *Circulation* **129**, 2100–2110 (2014).
- Elliott, A. D. et al. Association between physical activity and risk of incident arrhythmias in 402 406 individuals: evidence from the UK Biobank cohort. *Eur. Heart J.* **41**, 1479–1486 (2020).
- Qiu, S. & Xing, Z. Association between accelerometer-derived physical activity and incident cardiac arrest. *Europace* **25**, eua4353 (2023).
- Yang, L. et al. Association of accelerometer-derived circadian abnormalities and genetic risk with incidence of atrial fibrillation. *NPJ Digit Med* **6**, 31 (2023).
- Liebzeit, D., Phelan, C., Moon, C., Brown, R. & Bratzke, L. Rest-activity patterns in older adults with heart failure and healthy older adults. *J. Aging Phys. Act.* **25**, 116–122 (2017).
- Castela Forte, J. et al. Identifying and characterizing high-risk clusters in a heterogeneous ICU population with deep embedded clustering. *Sci. Rep.* **11**, 12109 (2021).

19. Rowlands, A. V. et al. Beyond cut points: accelerometer metrics that capture the physical activity profile. *Med Sci. Sports Exerc* **50**, 1323–1332 (2018).
20. Frodi, D. M. et al. Rationale and design of the SafeHeart study: Development and testing of a mHealth tool for the prediction of arrhythmic events and implantable cardioverter-defibrillator therapy. *Cardiovasc Digit Health J.* **2**, S11–S20 (2021).
21. von Elm, E. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J. Clin. Epidemiol.* **61**, 344–349 (2008).
22. Eslinger, D. W. et al. Validation of the GENEA accelerometer. *Med Sci. Sports Exerc* **43**, 1085–1093 (2011).
23. van Hees, V. T. et al. A novel, open access method to assess sleep duration using a wrist-worn accelerometer. *PLoS One* **10**, e0142533 (2015).
24. Spertus, J. A. & Jones, P. G. Development and Validation of a Short Version of the Kansas City Cardiomyopathy Questionnaire. *Circ. Cardiovasc Qual. Outcomes* **8**, 469–476 (2015).
25. Herdman, M. et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual. Life Res* **20**, 1727–1736 (2011).
26. Kingma, D. P. et al. Improved Variational Inference with Inverse Autoregressive Flow. Vol. 29 (eds Lee, D. et al.) (2016).
27. Higgins, I. et al. beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework. in *International Conference on Learning Representations* (2016).
28. Vouros, A., Langdell, S., Croucher, M. & Vasilaki, E. An empirical comparison between stochastic and deterministic centroid initialisation for K-means variations. *Mach. Learn.* **110**, 1975–2003 (2021).
29. Hennig, C. Cluster-wise assessment of cluster stability. *Computational Stat. Data Anal.* **52**, 258–271 (2007).
30. Pedregosa, F. et al. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
31. Scott, M. L. & Su-In, L. A Unified Approach to Interpreting Model Predictions. *ArXiv abs/1705.07874* (2017).
32. Tianqi, C. & Carlos, G. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).

Acknowledgements

This research was supported by the Horizon 2020 European Union funding programme for research and innovation (grant number: Eurostars project E! 113994- SafeHeart). This research is partly funded by the Amsterdam Cardiovascular Sciences and research programme Rubicon which is financed by the Dutch Research Council (NWO).

Author contributions

M.K., R.K., J.H.S., J.L., T.A., S.D. and F.T. contributed to the conception and design of the study. M.K., D.F., J.L. and F.T. collectively investigated the data and decided on the methodology to be used. M.K., F.T. and J.L. conducted the formal analyses. M.K. and F.T. drafted the original manuscript. M.K., D.F., J.L., T.A., P.K.J., N.R., H.T., J.H.S., S.D., R.K. and F.T. reviewed, edited, and agreed with the final version of the manuscript. M.K., J.L., D.F. and F.T. accessed and verified the underlying data.

Competing interests

The authors declare no competing non-financial interests but the following competing financial interests. R.E.K. reports consultancy fees and research grants from Boston Scientific, Medtronic and Abbott. and has stock options from AtaCor Medical Inc. F.V.Y.T. declares grants or contracts from the Dutch Research Council (NWO) and Amsterdam Cardiovascular Sciences, and received payment or honoraria from Boston Scientific and Abbott (paid to the institution). S.Z.D. reports consultancy fees and research grants from Acesion Pharma and Cortrium, and has received payment or honoraria from Bristol, Myers Squibb, Pfizer and Bayer. P.K.J. reports consultancy fees and research grants from Abbott and Medtronic, payment or honoraria from Abbott and Medtronic and support for attending meetings and/or travel from Abbott and Medtronic. J.H.S. reports grants or contracts from Medtronic (paid to institution), payment or honoraria from Medtronic, support for attending meetings and/or travel from Abbott and Medtronic, participation on Medtronic Advisory Board, and stocks or stock options from Vital Beats. J.L. reports having stock or stock option from Activinsights Ltd. T.O.A. reports having stock or stock options from Vital Beats. D.M.F. reports financial support for attending meetings and/or travel from Boston Scientific. The authors M.Z.H.K., N.R. and H.L.T. declare no competing financial or non-financial interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-024-01247-w>.

Correspondence and requests for materials should be addressed to Fleur V. Y. Tjong.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024