
Supplementary information

Arrhythmic sudden death survival prediction using deep learning analysis of scarring in the heart

In the format provided by the authors and unedited

| | Internal | | External | |
|--|--------------------|--------------------|--------------------|--------------------|
| | SSCAR | SSCAR (CMR Only) | SSCAR | SSCAR (CMR Only) |
| c-index | 0.83 (0.81 – 0.85) | 0.70 (0.67 – 0.72) | 0.74 (0.71 – 0.77) | 0.63 (0.59 – 0.66) |
| $\overline{\text{Bs}}$ | 0.12 (0.12 – 0.13) | 0.17 (0.17 – 0.18) | 0.14 (0.14 – 0.15) | 0.19 (0.19 – 0.20) |
| BA | 0.77 (0.73 – 0.81) | 0.60 (0.55 – 0.65) | 0.73 (0.69 – 0.77) | 0.63 (0.58 – 0.68) |
| F-score | 0.87 (0.84 – 0.89) | 0.81 (0.77 – 0.85) | 0.65 (0.59 – 0.70) | 0.60 (0.55 – 0.65) |

Supplementary Table 1: SSCAR Detailed Performance. Concordance index (c-index), integrated Brier score ($\overline{\text{Bs}}$), balanced accuracy (BA) and F-score are shown for the entire model — SSCAR — and for the CMR sub-network using on images only — SSCAR (CMR Only). All performance measures are calculated at $\tau = 10$ years. The numbers for the internal cohort (Internal) are averages over 100 cross-validation splits and values on the external data set (External) represent a single evaluation on the entire set. In parentheses, approximate 95% confidence intervals are shown (see Methods for details).

| Parameter | CMR [Choice] | Covariate [Choice] |
|-------------------------------------|---|---|
| Epochs (maximum) | 2000 | 2000 |
| Learning rate (initial) | .01 | .01 |
| Batch size | 32 | 32 |
| Steps per epoch | 20 | 20 |
| Activation function | ReLU | ReLU |
| Convolutional kernel size | $3 \times 3 \times 3$ | N/A |
| Network depth | {2, 4} [2] | {1, 2, 3, 4, 5} [3] |
| Risk categories | {1, 2, 3} [2] | {1, 2, 3} [1] |
| Dropout | Uniform(0, 0.5) [0.04] | Uniform(0, 0.5) [.22] |
| l1 (bias and kernel) regularization | LogUniform(10^{-7} , 10^{-3}) [4.02×10^{-5}] | LogUniform(10^{-7} , 10^{-3}) [1.41×10^{-5}] |
| l2 (bias and kernel) regularization | LogUniform(10^{-6} , 10^{-2}) [3.25×10^{-6}] | LogUniform(10^{-6} , 10^{-2}) [8.27×10^{-5}] |
| Convolutions per sampling level | {1, 2} [1] | N/A |
| Convolutional filters (initial) | {8, 12, 16, 20, 24} [24] | N/A |
| Reconstruction loss weight | Uniform(0, 10) [.06] | N/A |
| Latent representation dimension | Uniform{10, 30} [22] | N/A |
| Units per dense layer | N/A | Uniform{4, 24} [8] |

Supplementary Table 2: Hyperparameter Tuning. The hyperparameters used to define network architectures and training guidelines (left column) were optimized by randomly sampling using the Parzen windows algorithm from the search spaces defined for the CMR (middle column) and covariate sub-networks (right column). Braces denote equally probable sampling from the set. “Uniform” refers to uniformly distributed either continuously (parentheses) or discretely (braces). “LogUniform” refers to uniformly distributed exponent. No qualification means the hyperparameter was fixed. Network depth is the number of convolution/downsample steps for the CMR sub-network and the number of dense layers for the covariate sub-network. Risk categories refers to the number of strata used to divide before the final layer (see details in Methods). The latent representation dimension is the number of units of the dense layer immediately following the last convolution/downsample step. The hyperparameter search was stopped after 100 and 300 iterations for the CMR and covariate sub-networks, respectively. Choices for the final model are displayed in bold in square brackets.