

Supplementary Information

Stroke Onset Time Estimation from NCCT with Censoring-aware Learning and Robustness to Infarct Segmentation Uncertainty

Linda Vorberg, Leonhard Rist, Hendrik Ditt, Andreas Maier, and Oliver Taubmann

Contents. This document contains additional tables referenced from the main manuscript.

Supplementary Results

Supplementary Table S1. SVR-based ablation of strategies for incorporating last-known-well-time (LKWT) cases into onset-time regression on the external test cohort. All models use the same five radiomics features and RBF kernel. “Hybrid censoring-aware loss” corresponds to $\text{SVR}_{L_{\alpha=0.8}}$, and “Explicit censoring constraints” to SVR_{RR} in the main manuscript. Performance is reported for all 32 patients with documented onset time, retaining the predefined 0h estimate when no qualifying NCCT infarct was identified. Values are MedAE, MAE, and RMSE in hours with 95% BCa bootstrap confidence intervals from 10,000 patient-level resamples. The best point estimate per metric is shown in bold.

Training strategy for SVR	n_{train}	n_{test}	MedAE [h]	MAE [h]	RMSE [h]
Known-onset only	119	32	1.23 [1.05, 1.48]	2.42 [1.27, 4.97]	5.15 [1.41, 9.34]
LKWT as Surrogate	142	32	1.31 [1.18, 1.58]	2.67 [1.56, 5.23]	5.27 [1.85, 9.38]
Stochastic LKWT Inclusion	142	32	1.62 [1.23, 1.99]	2.62 [1.48, 5.12]	5.23 [1.69, 9.30]
Hybrid censoring-aware loss	142	32	1.16 [0.66, 1.52]	2.43 [1.29, 5.02]	5.22 [1.71, 9.37]
Explicit censoring constraints	142	32	1.12 [0.91, 1.42]	2.39 [1.24, 4.95]	5.15 [1.44, 9.35]

n_{train} , number of training cases; n_{test} , number of evaluated external test cases; MedAE, median absolute error; MAE, mean absolute error; RMSE, root mean squared error; LKWT, last-known-well time.

Supplementary Table S2. Onset-time regression errors across the four experimental settings on the external test set ($n = 32$), retaining the predefined 0h estimate for cases without a qualifying NCCT infarct. Values are point estimates with 95% BCa bootstrap confidence intervals from 10,000 patient-level resamples. The lowest point estimate within each panel is shown in bold. Models marked with † do not depend on LKWT target assignment; their identical values are repeated across the corresponding panels for completeness.

Model	MedAE [h] (95% CI)	MAE [h] (95% CI)	RMSE [h] (95% CI)
A. LKWT as surrogate target, without infarct contour perturbation			
LR	2.53 [1.57, 2.92]	3.60 [2.49, 5.94]	5.74 [3.01, 9.56]
RF	1.88 [1.31, 2.57]	3.10 [1.99, 5.59]	5.46 [2.38, 9.47]
SVR	1.31 [1.18, 1.58]	2.67 [1.56, 5.23]	5.27 [1.85, 9.38]
XGB	1.42 [1.17, 2.42]	3.00 [1.85, 5.46]	5.50 [2.48, 9.45]
EBM	1.64 [1.23, 2.14]	2.80 [1.68, 5.34]	5.32 [2.00, 9.43]
MLP $_{L_{\alpha=0.0}}$	1.25 [1.00, 1.36]	2.50 [1.37, 5.08]	5.23 [1.72, 9.39]
MLP $_{L_{\alpha=0.7}}$	1.21 [1.01, 1.38]	2.29 [1.14, 4.84]	5.12 [1.31, 9.32]

Continued on next page

Table S2 continued

Model	MedAE [h] (95% CI)	MAE [h] (95% CI)	RMSE [h] (95% CI)
SVR $_{L_{\alpha}=0.8}$	1.16 [0.66, 1.52]	2.43 [1.29, 5.02]	5.22 [1.71, 9.37]
SVR $_{RR}^{\dagger}$	1.12 [0.91, 1.42]	2.39 [1.24, 4.95]	5.15 [1.44, 9.35]
Mean-HU-ratio-LR	2.99 [1.47, 3.84]	3.85 [2.73, 6.13]	5.89 [3.30, 9.66]
NWU †	1.37 [1.11, 1.90]	2.64 [1.52, 5.18]	5.25 [1.79, 9.38]
B. LKWT-based target sampling, without infarct contour perturbation			
LR	2.72 [1.68, 3.43]	3.71 [2.63, 6.05]	5.75 [3.01, 9.56]
RF	2.43 [1.50, 3.40]	3.50 [2.41, 5.92]	5.63 [2.77, 9.55]
SVR	1.62 [1.23, 1.99]	2.62 [1.48, 5.12]	5.23 [1.69, 9.30]
XGB	1.70 [1.23, 2.25]	2.84 [1.72, 5.38]	5.33 [2.00, 9.37]
EBM	2.08 [1.55, 3.14]	3.48 [2.39, 5.85]	5.62 [2.75, 9.50]
$MLP_{0.0}$	1.38 [1.22, 2.12]	2.65 [1.52, 5.22]	5.25 [1.79, 9.35]
$MLP_{0.7}$	1.28 [0.98, 1.68]	2.44 [1.29, 5.01]	5.16 [1.47, 9.35]
$SVR_{0.8}$	1.24 [0.73, 1.68]	2.64 [1.51, 5.16]	5.30 [1.93, 9.36]
SVR $_{RR}^{\dagger}$	1.12 [0.91, 1.42]	2.39 [1.24, 4.95]	5.15 [1.44, 9.35]
Mean-HU-ratio-LR	2.42 [1.38, 2.96]	3.22 [2.12, 5.71]	5.50 [2.47, 9.46]
NWU †	1.37 [1.11, 1.90]	2.64 [1.52, 5.18]	5.25 [1.79, 9.38]
C. LKWT as surrogate target, with infarct contour perturbation			
LR	1.99 [1.51, 3.02]	3.40 [2.30, 5.87]	5.62 [2.76, 9.46]
RF	2.21 [1.41, 2.84]	3.33 [2.23, 5.86]	5.58 [2.66, 9.52]
SVR	1.48 [1.28, 1.81]	2.66 [1.54, 5.23]	5.24 [1.74, 9.38]
XGB	1.90 [1.22, 2.41]	3.12 [1.99, 5.64]	5.55 [2.59, 9.47]
EBM	1.90 [1.37, 2.43]	3.02 [1.90, 5.53]	5.41 [2.23, 9.45]
$MLP_{0.0}$	1.47 [1.21, 1.80]	2.74 [1.62, 5.28]	5.28 [1.89, 9.39]
$MLP_{0.7}$	1.25 [1.15, 1.45]	2.38 [1.23, 4.93]	5.14 [1.36, 9.33]
$SVR_{0.8}$	0.90 [0.50, 1.38]	2.30 [1.15, 4.90]	5.18 [1.56, 9.34]
SVR $_{RR}^{\dagger}$	1.24 [1.00, 1.50]	2.43 [1.28, 4.96]	5.15 [1.43, 9.31]
Mean-HU-ratio-LR	2.62 [1.32, 4.05]	3.75 [2.62, 6.09]	5.85 [3.24, 9.64]
NWU †	1.53 [1.23, 1.77]	2.62 [1.50, 5.20]	5.25 [1.76, 9.38]
D. LKWT-based target sampling, with infarct contour perturbation			
LR	2.29 [1.44, 2.93]	3.49 [2.41, 5.87]	5.63 [2.77, 9.52]
RF	2.43 [1.52, 3.06]	3.58 [2.49, 6.00]	5.69 [2.88, 9.54]
SVR	1.58 [1.23, 1.95]	2.59 [1.45, 5.12]	5.21 [1.64, 9.33]
XGB	1.49 [1.24, 2.13]	2.89 [1.78, 5.45]	5.37 [2.13, 9.41]
EBM	2.02 [1.50, 2.84]	3.39 [2.30, 5.79]	5.57 [2.63, 9.52]
$MLP_{0.0}$	1.40 [1.19, 2.09]	2.65 [1.53, 5.23]	5.25 [1.76, 9.35]
$MLP_{0.7}$	1.28 [0.98, 1.60]	2.41 [1.27, 4.97]	5.15 [1.43, 9.34]
$SVR_{0.8}$	1.24 [0.88, 1.69]	2.65 [1.53, 5.18]	5.29 [1.92, 9.36]
SVR $_{RR}^{\dagger}$	1.24 [1.00, 1.50]	2.43 [1.28, 4.96]	5.15 [1.43, 9.31]
Mean-HU-ratio-LR	2.21 [1.26, 2.95]	3.15 [2.05, 5.66]	5.49 [2.44, 9.47]
NWU †	1.53 [1.23, 1.77]	2.62 [1.50, 5.20]	5.25 [1.76, 9.38]

Supplementary Table S3. Significant pairwise model comparisons in the primary experiment (LKWT as surrogate, without contour perturbation). Two-sided Wilcoxon signed-rank tests compared patient-level absolute errors among the 22 cases receiving feature-based predictions. The 10 cases receiving the predefined 0h estimate were identical across models and therefore provided zero paired differences. Holm correction was applied across all 55 pairwise comparisons; only adjusted $p < 0.05$ are shown.

Model with lower paired absolute errors	Models with higher paired absolute errors (p_{Holm})
EBM	LR (0.012), Mean-HU-Ratio-LR (0.019)
MLP _{$L_{\alpha=0.0}$}	LR (0.002), Mean-HU-Ratio-LR (0.012)
MLP _{$L_{\alpha=0.7}$}	LR (0.009), Mean-HU-Ratio-LR (0.006)
NWU	LR (0.022), Mean-HU-Ratio-LR (0.002)
SVR _{RR}	LR (0.012), Mean-HU-Ratio-LR (0.009)
SVR	LR (0.002), Mean-HU-Ratio-LR (0.006)

Models not listed showed no Holm-significant pairwise differences. The direction was determined from the median paired difference in absolute error.

Supplementary Table S4. Onset-time regression errors across the four experimental settings in the external test set for the **n=22** cases with a qualifying infarct detected in NCCT, excluding 10 cases with the predefined 0h estimate. Values are MedAE, MAE, and RMSE in hours with 95% BCa bootstrap confidence intervals from 10,000 patient-level resamples. The lowest point estimate per metric within each panel is shown in bold. Models marked with † are unaffected by LKWT target assignment and are repeated across the corresponding panels.

Model	MedAE [h] (95% CI)	MAE [h] (95% CI)	RMSE [h] (95% CI)
A. LKWT as surrogate target, without infarct contour perturbation			
LR	2.79 [2.33, 3.47]	2.98 [2.38, 3.68]	3.36 [2.77, 4.11]
RF	2.28 [1.50, 2.96]	2.26 [1.74, 2.83]	2.61 [2.14, 3.27]
SVR	1.31 [1.00, 1.70]	1.64 [1.25, 2.16]	1.96 [1.51, 2.56]
XGB	1.53 [0.57, 2.96]	2.11 [1.45, 2.88]	2.72 [2.06, 3.56]
EBM	1.82 [1.16, 2.41]	1.83 [1.38, 2.33]	2.16 [1.77, 2.75]
MLP _{0.0}	1.25 [0.87, 1.37]	1.39 [1.02, 2.03]	1.80 [1.25, 2.69]
MLP _{0.7}	1.07 [0.62, 1.38]	1.09 [0.83, 1.38]	1.28 [1.03, 1.59]
SVR _{0.8}	0.75 [0.51, 1.63]	1.29 [0.87, 1.90]	1.76 [1.24, 2.49]
SVR _{RR} [†]	1.00 [0.70, 1.37]	1.23 [0.96, 1.60]	1.44 [1.13, 1.91]
Mean-HU-ratio-LR	3.34 [2.75, 4.34]	3.36 [2.70, 4.02]	3.72 [3.17, 4.44]
NWU [†]	1.49 [0.88, 1.96]	1.59 [1.22, 2.08]	1.89 [1.50, 2.53]
B. LKWT-based target sampling, without infarct contour perturbation			
LR	3.13 [2.56, 3.74]	3.15 [2.63, 3.66]	3.39 [2.92, 3.91]
RF	2.79 [2.15, 3.57]	2.84 [2.31, 3.33]	3.09 [2.62, 3.57]
SVR	1.78 [1.09, 2.12]	1.56 [1.17, 1.89]	1.78 [1.48, 2.05]
XGB	1.92 [0.98, 2.34]	1.88 [1.43, 2.33]	2.17 [1.76, 2.66]
EBM	2.81 [1.67, 3.58]	2.82 [2.34, 3.36]	3.07 [2.59, 3.65]
MLP _{0.0}	1.47 [0.99, 2.38]	1.61 [1.20, 2.03]	1.90 [1.54, 2.26]
MLP _{0.7}	1.22 [0.84, 1.71]	1.30 [1.02, 1.64]	1.49 [1.22, 1.89]
SVR _{0.8}	0.96 [0.61, 2.16]	1.59 [1.11, 2.21]	2.06 [1.54, 2.71]
SVR _{RR} [†]	1.00 [0.70, 1.37]	1.23 [0.96, 1.60]	1.44 [1.13, 1.91]

Continued on next page

Table S4 continued

Model	MedAE [h] (95% CI)	MAE [h] (95% CI)	RMSE [h] (95% CI)
Mean-HU-ratio-LR	2.60 [1.98, 3.19]	2.44 [1.90, 2.94]	2.74 [2.33, 3.21]
NWU [†]	1.49 [0.88, 1.96]	1.59 [1.22, 2.08]	1.89 [1.50, 2.53]
C. LKWT as surrogate target, with infarct contour perturbation			
LR	2.38 [1.69, 3.55]	2.69 [2.11, 3.35]	3.07 [2.51, 3.73]
RF	2.51 [1.58, 3.09]	2.59 [2.03, 3.22]	2.95 [2.42, 3.66]
SVR	1.53 [1.33, 1.94]	1.63 [1.30, 2.02]	1.84 [1.55, 2.36]
XGB	2.37 [1.04, 2.62]	2.29 [1.66, 3.10]	2.86 [2.20, 3.73]
EBM	2.22 [1.61, 2.53]	2.14 [1.68, 2.66]	2.44 [2.02, 3.03]
$MLP_{0.0}$	1.52 [1.07, 2.00]	1.73 [1.35, 2.22]	2.02 [1.60, 2.57]
$MLP_{0.7}$	1.22 [0.89, 1.45]	1.21 [0.96, 1.47]	1.35 [1.14, 1.64]
$SVR_{0.8}$	0.53 [0.33, 1.12]	1.10 [0.70, 1.67]	1.58 [1.09, 2.19]
$SVR_{RR}^{\dagger}$	1.11 [0.83, 1.49]	1.28 [1.03, 1.59]	1.45 [1.17, 1.80]
Mean-HU-ratio-LR	3.55 [2.38, 4.22]	3.21 [2.50, 3.94]	3.65 [3.07, 4.44]
NWU [†]	1.59 [1.19, 1.86]	1.56 [1.19, 2.04]	1.86 [1.51, 2.61]
D. LKWT-based target sampling, with infarct contour perturbation			
LR	2.76 [2.23, 3.28]	2.83 [2.34, 3.38]	3.09 [2.62, 3.67]
RF	2.91 [2.30, 3.61]	2.96 [2.42, 3.51]	3.24 [2.75, 3.78]
SVR	1.75 [1.09, 2.10]	1.51 [1.15, 1.82]	1.71 [1.44, 1.96]
XGB	1.68 [1.24, 2.39]	1.96 [1.53, 2.59]	2.32 [1.83, 3.15]
EBM	2.69 [1.66, 3.12]	2.68 [2.23, 3.21]	2.93 [2.46, 3.51]
$MLP_{0.0}$	1.55 [1.08, 2.29]	1.61 [1.22, 2.00]	1.87 [1.52, 2.20]
$MLP_{0.7}$	1.33 [0.76, 1.66]	1.26 [0.98, 1.57]	1.44 [1.19, 1.79]
$SVR_{0.8}$	1.02 [0.64, 2.17]	1.61 [1.14, 2.21]	2.05 [1.54, 2.67]
$SVR_{RR}^{\dagger}$	1.11 [0.83, 1.49]	1.28 [1.03, 1.59]	1.45 [1.17, 1.80]
Mean-HU-ratio-LR	2.87 [1.97, 3.03]	2.34 [1.76, 2.87]	2.70 [2.27, 3.20]
NWU [†]	1.59 [1.19, 1.86]	1.56 [1.19, 2.04]	1.86 [1.51, 2.61]