Supplementary Information

Prediction of ten intraoperative physiologic derangements using routine monitoring and drug administration data

Jake Stephen, Mitchell Ng, Kenneth Ng, Jonathan W. Pillow and Armita Kazemi

Corresponding author: Jake Stephen (js6781@princeton.edu)

Contents

Supplementary Table 1. Discrimination (AUROC) for all ten targets at all three horizons, both cohorts
Supplementary Table 2. Area under the precision-recall curve for all ten targets at all three horizons, both cohorts
Supplementary Table 3. Per-target operating-point detail at the standard budget and 5-min horizon
Supplementary Table 4. Warning time by target and horizon, both cohorts
Supplementary Table 5. Calibration by target and horizon, both cohorts
Supplementary Table 6. Mean stable AUROC at the 5-min horizon by subgroup, both cohorts
Supplementary Table 7. Illustrative three-target configuration at the conservative budget and 10-min horizon

Supplementary Table 1 | Discrimination for all ten targets at all three horizons, both cohorts

a) Stable AUROC

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	0.900	0.872	0.830	0.873	0.850	0.816
Severe hypotension	0.882	0.858	0.819	0.876	0.858	0.832
Relative hypotension	0.900	0.885	0.860	0.865	0.837	0.789
Bradycardia	0.931	0.921	0.907	0.893	0.880	0.862
Tachycardia	0.864	0.845	0.811	0.884	0.862	0.823
Hypoxemia	0.948	0.930	0.868	0.837	0.811	0.767
Severe hypoxemia	0.960	0.944	0.886	0.841	0.817	0.772
Hypocapnia	0.854	0.832	0.789	0.801	0.784	0.753
Hypercapnia	0.967	0.944	0.879	0.903	0.883	0.840
Hypothermia	0.847	0.836	0.811	0.853	0.843	0.825
Mean	**0.905**	**0.887**	**0.846**	**0.863**	**0.843**	**0.808**

AUROC = area under the receiver operating characteristic curve; Dev. = development test partition; Ext. = external cohort. Stable AUROC is computed only over windows in which the patient was not already breaching the threshold for that target at the time of prediction.

b) Standard AUROC

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	0.945	0.921	0.881	0.944	0.921	0.884
Severe hypotension	0.926	0.896	0.851	0.935	0.910	0.873
Relative hypotension	0.973	0.962	0.943	0.965	0.949	0.916
Bradycardia	0.980	0.974	0.962	0.969	0.959	0.942
Tachycardia	0.954	0.932	0.894	0.965	0.946	0.910
Hypoxemia	0.978	0.962	0.909	0.908	0.872	0.816
Severe hypoxemia	0.981	0.967	0.916	0.908	0.872	0.815
Hypocapnia	0.927	0.899	0.847	0.907	0.887	0.851
Hypercapnia	0.978	0.958	0.896	0.946	0.930	0.894

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypothermia	0.977	0.962	0.931	0.977	0.967	0.947
Mean	**0.962**	**0.943**	**0.903**	**0.942**	**0.921**	**0.885**

Standard AUROC is computed over all evaluable windows.

Supplementary Table 2 | Area under the precision-recall curve for all ten targets at all three horizons, both cohorts

a) Stable area under the precision-recall curve

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	0.373	0.414	0.477	0.242	0.287	0.342
Severe hypotension	0.199	0.224	0.261	0.126	0.155	0.193
Relative hypotension	0.421	0.496	0.576	0.297	0.347	0.389
Bradycardia	0.450	0.514	0.582	0.356	0.421	0.491
Tachycardia	0.236	0.288	0.348	0.167	0.208	0.252
Hypoxemia	0.250	0.315	0.349	0.037	0.041	0.047
Severe hypoxemia	0.200	0.256	0.292	0.027	0.031	0.032
Hypocapnia	0.214	0.263	0.320	0.192	0.248	0.313
Hypercapnia	0.080	0.082	0.073	0.097	0.120	0.152
Hypothermia	0.107	0.159	0.228	0.066	0.097	0.142
Mean	**0.253**	**0.301**	**0.350**	**0.161**	**0.196**	**0.235**

Episode prevalence differs by target and cohort, so these values are not comparable across rows; they are reported to indicate the precision attainable on the pre-event index set.

b) Standard area under the precision-recall curve

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	0.740	0.727	0.719	0.733	0.707	0.675
Severe hypotension	0.491	0.451	0.423	0.526	0.482	0.431
Relative hypotension	0.938	0.929	0.917	0.911	0.896	0.864
Bradycardia	0.931	0.927	0.917	0.912	0.905	0.890

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Tachycardia	0.808	0.772	0.728	0.809	0.772	0.714
Hypoxemia	0.748	0.708	0.619	0.345	0.278	0.205
Severe hypoxemia	0.663	0.619	0.532	0.302	0.238	0.167
Hypocapnia	0.655	0.625	0.584	0.733	0.718	0.699
Hypercapnia	0.264	0.229	0.171	0.475	0.466	0.451
Hypothermia	0.919	0.885	0.827	0.907	0.882	0.836
Mean	**0.716**	**0.687**	**0.644**	**0.665**	**0.634**	**0.593**

Computed over all evaluable windows.

Supplementary Table 3 | Per-target operating-point detail at the standard budget and 5-min horizon

a) Development test partition

Target	Threshold	Sens.	Spec.	PPV	NPV	Median lead, min [25th, 75th]	≥3 min, %
Hypotension	0.352	0.423	0.955	0.488	0.942	3.0 [2.0, 5.0]	53.8
Severe hypotension	0.051	0.597	0.906	0.137	0.989	4.0 [2.0, 6.0]	68.2
Relative hypotension	0.418	0.449	0.958	0.542	0.941	3.0 [2.0, 5.0]	54.4
Bradycardia	0.315	0.579	0.957	0.506	0.968	3.0 [2.0, 5.0]	56.1
Tachycardia	0.110	0.515	0.935	0.265	0.977	4.0 [2.0, 6.0]	63.7
Hypoxemia	0.037	0.796	0.933	0.191	0.996	4.0 [2.0, 7.0]	72.2
Severe hypoxemia	0.021	0.867	0.909	0.118	0.998	4.0 [2.0, 7.0]	73.0
Hypocapnia	0.165	0.470	0.933	0.288	0.968	4.0 [2.0, 6.0]	64.0
Hypercapnia	0.008	0.818	0.906	0.023	0.999	4.0 [2.0, 6.5]	69.9
Hypothermia	0.070	0.577	0.926	0.163	0.989	5.0 [3.0, 7.0]	75.4
Mean	**—**	**0.609**	**0.932**	**0.272**	**0.977**	**3.8**	**65.1**

NPV = negative predictive value; PPV = positive predictive value; Sens. = sensitivity; Spec. = specificity. Lead-time statistics are conditional on an alarm having been raised.

b) External cohort

Target	Threshold	Sens.	Spec.	PPV	NPV	Median lead, min [25th, 75th]	≥3 min, %
Hypotension	0.230	0.429	0.940	0.313	0.962	3.0 [2.0, 6.0]	58.6
Severe hypotension	0.047	0.541	0.926	0.117	0.991	4.0 [2.0, 6.0]	67.6
Relative hypotension	0.316	0.382	0.944	0.387	0.943	3.0 [2.0, 5.0]	56.0
Bradycardia	0.360	0.465	0.953	0.465	0.953	3.0 [2.0, 5.0]	54.4
Tachycardia	0.117	0.542	0.933	0.202	0.985	4.0 [2.0, 6.0]	62.8
Hypoxemia	0.019	0.473	0.921	0.038	0.996	4.0 [2.0, 6.0]	66.0
Severe hypoxemia	0.006	0.510	0.914	0.015	0.999	4.0 [2.0, 6.0]	68.4
Hypocapnia	0.303	0.326	0.938	0.297	0.946	4.0 [2.0, 6.0]	63.5
Hypercapnia	0.086	0.547	0.924	0.101	0.992	4.0 [2.0, 6.0]	64.6
Hypothermia	0.114	0.490	0.927	0.110	0.990	5.0 [3.0, 7.0]	75.7
Mean	**—**	**0.471**	**0.932**	**0.204**	**0.976**	**3.8**	**63.8**

Thresholds were re-selected on this cohort so that the same alarm budget was held fixed.

Supplementary Table 4 | Warning time by target and horizon, both cohorts

a) Median lead time, min [25th, 75th percentile], at the standard budget

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	2 [1, 3]	3 [2, 5]	4 [2, 8]	2 [1, 4]	3 [2, 6]	5 [2, 10]
Severe hypotension	3 [2, 4]	4 [2, 6]	6 [3, 11]	3 [2, 4]	4 [2, 6]	7 [3, 12]
Relative hypotension	2 [1, 4]	3 [2, 5]	4 [2, 8]	2 [1, 4]	3 [2, 5]	4 [2, 10]
Bradycardia	2 [1, 4]	3 [2, 5]	4 [2, 8]	2 [1, 4]	3 [2, 5]	4 [2, 8]
Tachycardia	2 [1, 4]	4 [2, 6]	6 [2, 11]	2 [1, 4]	4 [2, 6]	5 [2, 10]
Hypoxemia	3 [2, 4]	4 [2, 7]	7 [4, 12]	2 [1, 4]	4 [2, 6]	6 [3, 12]
Severe hypoxemia	3 [2, 4]	4 [2, 7]	8 [4, 12]	3 [2, 4]	4 [2, 6]	7 [3, 12]
Hypocapnia	2 [1, 4]	4 [2, 6]	5 [2, 10]	2 [1, 4]	4 [2, 6]	5 [2, 10]
Hypercapnia	3 [2, 4]	4 [2, 6.5]	6 [3, 11]	2 [1, 4]	4 [2, 6]	6 [3, 11]
Hypothermia	3 [2, 4]	5 [3, 7]	9 [4, 14]	3 [2, 4]	5 [3, 7]	9 [5, 14]

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Mean	**2.5**	**3.8**	**5.9**	**2.3**	**3.8**	**5.8**

Lead-time statistics are conditional on an alarm having been raised.

b) Proportion of true-positive warnings delivering at least 3 min of lead, %

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	41.7	53.8	64.2	44.1	58.6	70.6
Severe hypotension	53.6	68.2	77.3	51.6	67.6	79.3
Relative hypotension	40.9	54.4	64.8	41.7	56.0	67.5
Bradycardia	43.1	56.1	65.4	40.8	54.4	63.9
Tachycardia	47.3	63.7	74.7	47.6	62.8	73.4
Hypoxemia	55.4	72.2	82.7	49.7	66.0	78.0
Severe hypoxemia	55.8	73.0	83.4	52.2	68.4	80.3
Hypocapnia	49.3	64.0	74.9	48.6	63.5	73.3
Hypercapnia	53.5	69.9	79.6	49.0	64.6	75.3
Hypothermia	57.8	75.4	86.6	58.3	75.7	87.2
Mean	**49.8**	**65.1**	**75.4**	**48.4**	**63.8**	**74.9**

The prespecified threshold of clinical meaningfulness was 3 min.

Supplementary Table 5 | Calibration by target and horizon, both cohorts

a) Brier score (lower is better)

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	0.054	0.072	0.105	0.044	0.057	0.077
Severe hypotension	0.016	0.024	0.040	0.014	0.019	0.029
Relative hypotension	0.050	0.062	0.082	0.055	0.067	0.086
Bradycardia	0.039	0.046	0.059	0.049	0.059	0.076
Tachycardia	0.029	0.039	0.060	0.025	0.033	0.048
Hypoxemia	0.012	0.017	0.030	0.005	0.008	0.014

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Severe hypoxemia	0.009	0.013	0.023	0.002	0.003	0.005
Hypocapnia	0.037	0.051	0.081	0.070	0.085	0.110
Hypercapnia	0.002	0.003	0.006	0.018	0.023	0.032
Hypothermia	0.018	0.025	0.040	0.024	0.030	0.041
Mean	**0.026**	**0.035**	**0.053**	**0.031**	**0.038**	**0.052**

Calibrators (isotonic regression) were fitted on the development validation partition and applied unchanged to the external cohort.

b) Expected calibration error (lower is better)

Target	Dev. 3 min	Dev. 5 min	Dev. 10 min	Ext. 3 min	Ext. 5 min	Ext. 10 min
Hypotension	0.0034	0.0051	0.0088	0.0023	0.0023	0.0019
Severe hypotension	0.0016	0.0024	0.0037	0.0014	0.0016	0.0024
Relative hypotension	0.0054	0.0047	0.0077	0.0019	0.0023	0.0029
Bradycardia	0.0037	0.0042	0.0041	0.0017	0.0020	0.0024
Tachycardia	0.0034	0.0040	0.0088	0.0033	0.0043	0.0057
Hypoxemia	0.0016	0.0021	0.0028	0.0003	0.0003	0.0006
Severe hypoxemia	0.0012	0.0017	0.0018	0.0002	0.0002	0.0002
Hypocapnia	0.0019	0.0026	0.0043	0.0040	0.0047	0.0061
Hypercapnia	0.0007	0.0010	0.0019	0.0009	0.0008	0.0018
Hypothermia	0.0017	0.0023	0.0036	0.0027	0.0030	0.0037
Mean	**0.0025**	**0.0030**	**0.0047**	**0.0019**	**0.0022**	**0.0028**

Supplementary Table 6 | Mean stable AUROC at the 5-min horizon by subgroup, both cohorts

Subgroup	Stratum	Development	External	Difference
Age	< 50 yr	0.892	0.852	0.040
Age	50 to 70 yr	0.881	0.840	0.041
Age	≥ 70 yr	0.865	0.821	0.044

Subgroup	Stratum	Development	External	Difference
Sex	Female	0.889	0.847	0.043
Sex	Male	0.884	0.839	0.045
Case duration	< 2 h	0.895	0.851	0.043
Case duration	≥ 2 h	0.871	0.829	0.042

Values are mean stable AUROC across the ten targets. Sex was prespecified as a reporting stratum; age and case duration were not prespecified and are reported as consistency checks rather than as a formal fairness evaluation.

Supplementary Table 7 | Illustrative three-target configuration at the conservative budget and 10-min horizon

Target	Sens., dev.	PPV, dev.	Sens., ext.	PPV, ext.	Median lead, min
Relative hypotension	0.21	0.75	0.20	0.53	4
Bradycardia	0.28	0.74	0.21	0.64	4
Hypotension	0.19	0.63	0.19	0.46	4 to 5
Mean	**0.22**	**0.71**	**0.20**	**0.54**	**4**

PPV = positive predictive value; Sens. = sensitivity. Together these three targets consume approximately three alarms per case-hour. The configuration was selected by inspecting test-set performance and is illustrative rather than validated; a prespecified configuration confirmed on held-out data would be required before any clinical claim.