

Table of contents

Supplemental Figure 1. Time frame of the study

Supplemental Figure 2. Development and validation process of machine learning models and integer-based score system (DPC score) to predict parathyroidectomy after kidney transplantation

Supplemental Figure 3. The threshold for DPC scores to detect individuals at high risk of post-transplant parathyroidectomy was determined at the point that maximized F1-score (balance between precision and recall) in train set (high risk ≥ 13 vs. low risk <13).

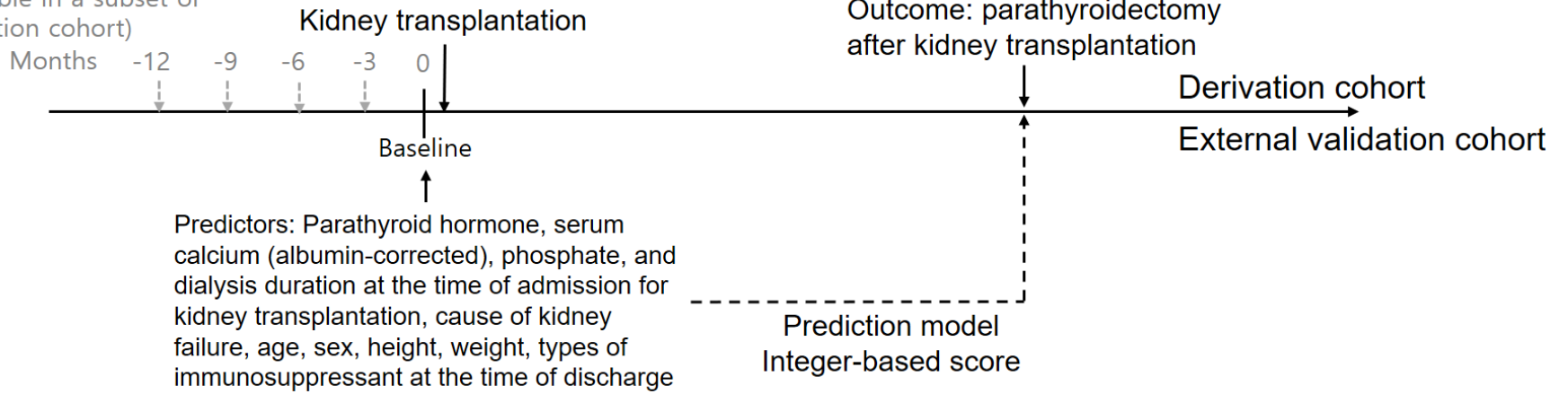
Supplemental Figure 4. Comparison of discriminatory ability for post-transplant parathyroidectomy between best-performing machine learning model (extreme gradient boosting model) and integer-based score (DPC score) in internal test set and external validation cohort

Supplemental Figure 5. The time-dependent trajectory of machine learning-derived integer-based score (DPC score) during 12 months prior to kidney transplantation in a subset of derivation cohort with measurements at two or more time points (at least three-month interval). Markers and capped spikes indicate median and interquartile range, respectively. *: $p < 0.05$ vs. no post-transplant parathyroidectomy group in a linear mixed model at each time point.

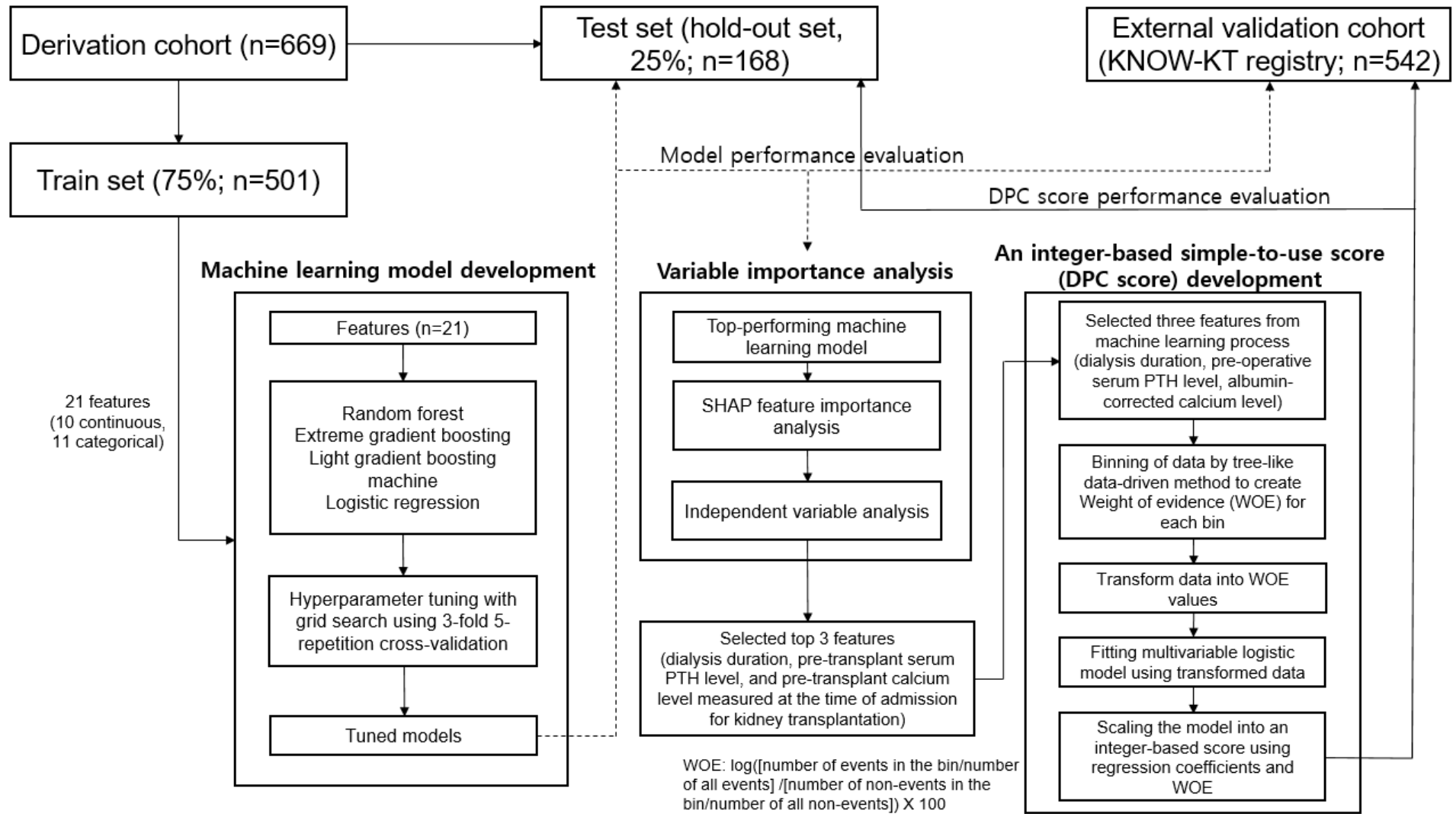
Supplemental Table 1. Model hyperparameters

Supplemental Table 2. Comparison of clinical characteristics of derivation and external validation cohort

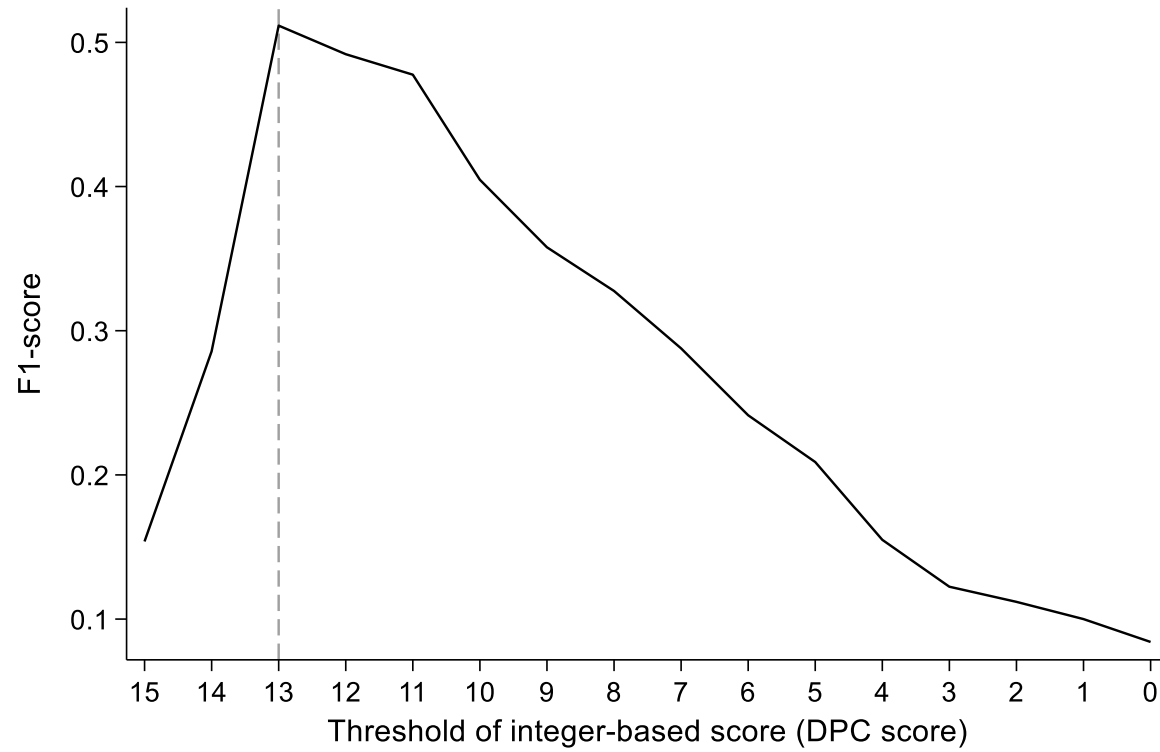
Serial pre-transplant
biochemical measurements
(available in a subset of
derivation cohort)



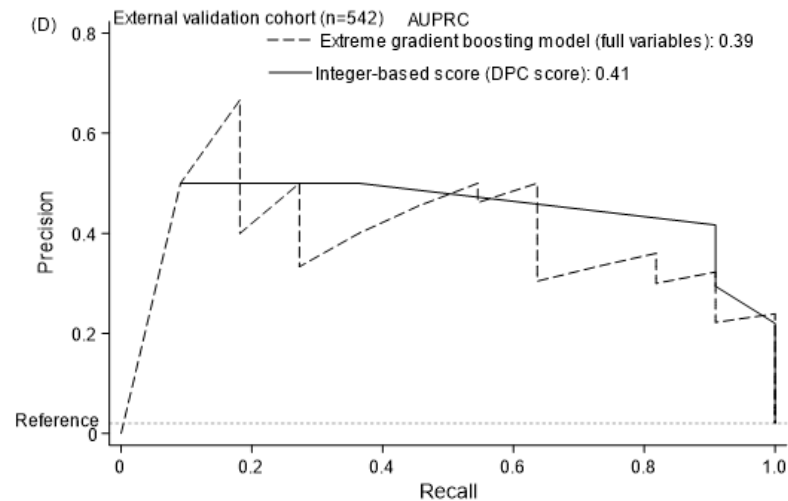
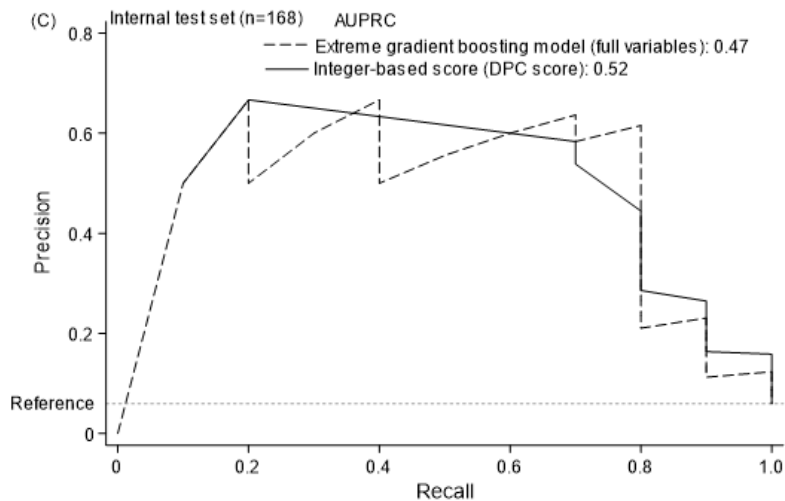
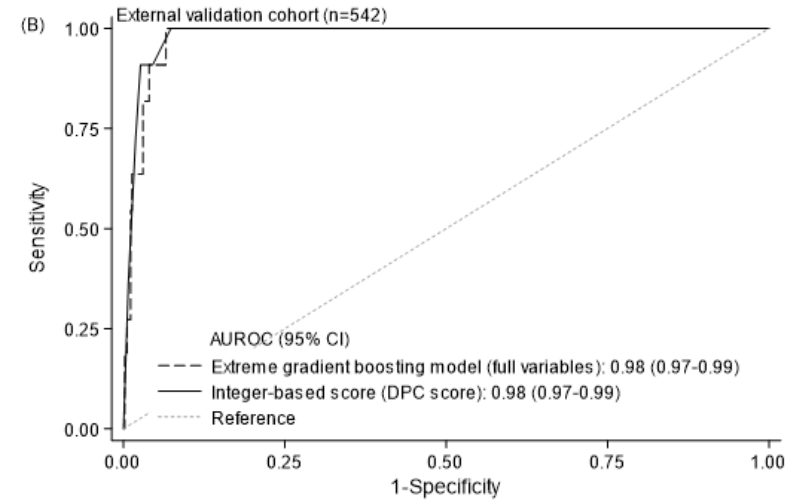
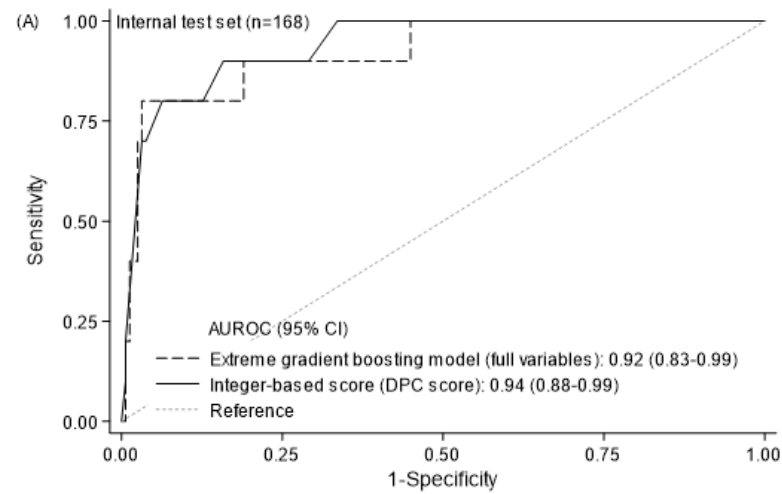
Supplemental Figure 1. Time frame of the study



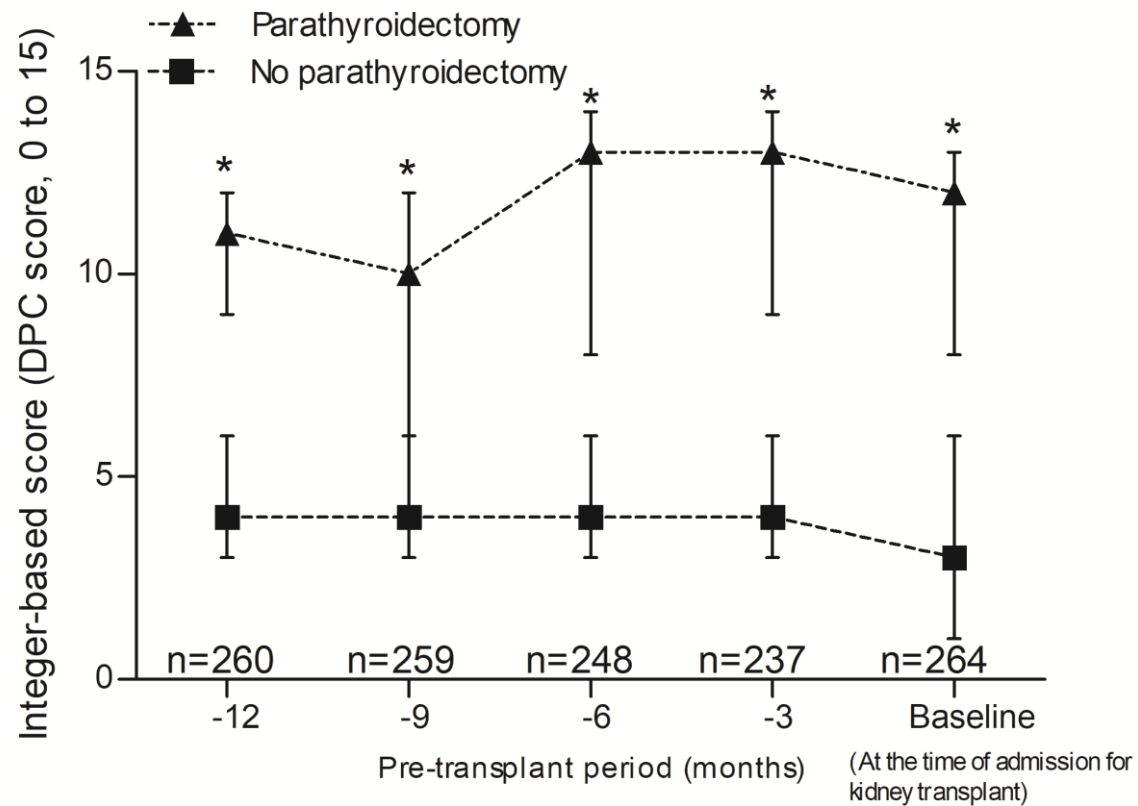
Supplemental Figure 2. Development and validation process of machine learning models and integer-based score system (DPC score) to predict parathyroidectomy after kidney transplantation



Supplemental Figure 3. The threshold for DPC scores to detect individuals at high risk of post-transplant parathyroidectomy was determined at the point that maximized F1-score (balance between precision and recall) in train set (high risk ≥ 13 vs. low risk <13).



Supplemental Figure 4. Comparison of discriminatory ability for post-transplant parathyroidectomy between best-performing machine learning model (extreme gradient boosting model) and integer-based score (DPC score) in internal test set and external validation cohort



Supplemental Figure 5. The time-dependent trajectory of machine learning-derived integer-based score (DPC score) during 12 months prior to kidney transplantation in a subset of derivation cohort (n=264/669, 39.4%) with measurements at two or more time points (at least three-month interval). Markers and capped spikes indicate median and interquartile range, respectively. *: p<0.05 vs. no post-transplant parathyroidectomy group in a linear mixed model at each time point.

Supplemental Table 1. Model hyperparameters

Models	Hyperparameters*
Random forest	Max_depth=24, min_samples_split=10, min_samples_leaf=2, max_features='log2'
Extreme gradient boosting	booster='gbtree', tree_method='exact', learning_rate=0.30, max_depth=5, max_leaves=255, subsample=1, colsample_bytree=0.8, gamma=2
Light gradient boosting machine	max_depth=-1, num_leaves=10, lambda_l2=1, min_data_in_leaf=100, reg_alpha=0.1
Logistic regression	C=0.1, solver='liblinear'

*Hyperparameters that were not stated were set to default values.

Supplemental Table 2. Comparison of clinical characteristics of derivation and external validation cohort

Variables	Derivation set (Severance hospital cohort) (n=699)	External validation set (KNOW-KT registry, a multi-center, prospective cohort) (n=542)	P value
Age, year	46 ± 11	46 ± 12	0.887
Women, n (%)	273 (41)	180 (33)	0.007
Body mass index, kg/m ²	22.4 ± 3.1	23.0 ± 3.5	0.003
Dialysis duration, months	8 [2-56]	5 [1-30]	<0.001
Calcium (albumin- corrected), mg/dL	8.4 ± 1.2	9.0 ± 1.0	<0.001
Phosphate, mg/dL	5.1 ± 1.3	5.1 ± 1.5	0.403
Albumin, g/dL	4.0 ± 0.5	4.0 ± 0.5	0.006
Parathyroid hormone, pg/mL	202 [108-354]	205 [107-323]	0.669
25-hydroxyvitamin D, ng/mL	11.6 [7.3-15.4]	10.6 [7.2-15.2]	0.171
Tacrolimus	612 (91)	522 (96)	0.001
Mycophenolate mofetil	597 (89)	484 (89)	0.973
Cyclosporine	135 (20)	16 (3)	<0.001
Cause of kidney failure			<0.001
Diabetic kidney disease	312 (47)	134 (25)	
Hypertensive kidney disease	153 (23)	103 (19)	
Glomerular disease	161 (24)	192 (35)	
Other causes	26 (4)	59 (11)	
Unspecified	17 (3)	54 (10)	