SUPPLEMENTAL DATA

Supplementary table 1. TRIPOD checklist for multivariable model development and validation.

Section/Topic	Item	Checklist item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	3
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	4-5
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	6
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	Describe eligibility criteria for participants.	5-6
	5c	Give details of treatments received, if relevant.	N/A
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	Report any actions to blind assessment of the outcome to be predicted.	7
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	7
Sample size	8	Explain how the study size was arrived at.	8
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	8-9
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8-9
	10c	For validation, describe how the predictions were calculated.	10
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
Risk groups	11	Provide details on how risk groups were created, if done.	N/A
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	6-7
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	10
Model	14a	Specify the number of participants and outcome events in each analysis.	10

development	14b	If done, report the unadjusted association between each candidate predictor and outcome.	10
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	10
	15b	Explain how to the use the prediction model.	10-11
Model performance	16	Report performance measures (with CIs) for the prediction model.	10-11
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	N/A
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	15
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	11-12
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	15-16
Implications	20	Discuss the potential clinical use of the model and implications for future research.	15
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	16
Funding	22	Give the source of funding and the role of the funders for the present study.	16

Supplementary table 2. The list of unusual infections according to our criteria, which were excluded from control datasets.

Diseases	Group
Aspergillosis	**Fungal and mycobacterial**
Blastomycosis*	
Coccidioidomycosis	
Cryptococcosis*	
Histoplasmosis*	
Mucormycosis*	
Nontuberculous mycobacterial infections	
Pneumocystosis*	
Tuberculosis*	
Acanthamoeba infections	**Parasitic**
Ascariasis	
Babesiosis	
Cysticercosis	
Filariasis	
Leishmaniasis	
Malariasis	
Naegleriasis	
Paragonimiasis	
Schistosomiasis	
Strongyloidiasis	
Toxocariasis	

Toxoplasmosis	
Trichinosis	
Trypanosomiasis	
Anaplasmosis	**Bacterial**
Bartonellosis	
Borrelia infections	
Coxiella burnetii infections	
Ehrlichiosis	
Legionellosis	
Leptospirosis	
Lymphogranuloma venereum	
Psittacosis	
Rickettsiosis	
Whipple's disease	
California encephalitis	**Viral**
Chikungunya	
Colorado tick fever	
Crimean-Congo hemorrhagic fever	
Dengue fever	
Ebola virus disease	
Hantavirus infections	
Japanese encephalitis	
Lymphocytic choriomeningitis	
Rift Valley fever	
Tick-borne relapsing fever	
Yellow fever	

* Infections that were focused on with this project.

Supplementary table 3. Variables evaluated for inclusion in the model and their definitions.

Variable name	Type	Definition	Missingness*	Included**	Similar domains	Stayed***
Age	Continuous	Patients' age, in years, determined from admission records	0	**Y**	**N**	**Y**
Sex	Categorical	Binary, determined from admission records: - Female - Male	0.1%	**Y**	**N**	**Y**
Race	Categorical	Four categories determined from admission records: - African American - Asian - White - Others	0.08%	**Y**	**N**	**Y**
Ethnicity	Categorical	Three categories determined from admission records: - Hispanic or Latino - Not Hispanic or Latino - Others, unknown, or not applicable	0.09%	**Y**	**N**	**N**
Quarter of admission	Categorical	Four levels according to the month of admission: - January to March - April to June - July to September - October to December	0	**Y**	**N**	**Y**
Admission location	Categorical	Four levels according to the specific Mayo Clinic Enterprise Hospital that the patient was admitted to: - Mayo Clinic Rochester - Mayo Clinic Florida - Mayo Clinic Arizona - Mayo Clinic Health System Hospitals, combined	0	**Y**	**N**	**Y**
Admission source	Categorical	Three levels according to the location before admission: - Another hospital or care facility - Outpatient or emergency department - Others or unknown	0.01%	**Y**	**Y, pre-admission location**	**Y**

Pre-hospital location home	Categorical	Binary, determined from admission records	27.3%	**Y**	**Y, pre-admission location**	
Transferred patient	Categorical	Binary, determined from admission records	27.3%	**Y**	**N**	**Y**
Country of residence	Categorical	Binary, according to the patients' residence address: - United States or Canada - Others	0.04%	**Y**	**Y, residency location**	**N**
World Health Organization region of residence	Categorical	Seven categories determined by the classification of patients' caountries (those other than United States or Canada) to World Health Organization Regions[41]: - African Region - Eastern Mediterranean Region - European Region - Region of the Americas, other than the US and Canada - South-East Asian Region - Western Pacific Region	2.00%	**Y**	**Y, residency location**	**N**
Rural-Urban Commuting Area codes	Categorical	Five categories determined by the classification of patient zip codes to Rural-Urban Commuting Area (RUCA) codes[42], which is used to categorize regions into 1 (most urban) to 10 (most rural)[43]. For the sake of simplicity, RUCA codes were grouped into four categories [28]: - Metropolitan area - Micropolitan area - Small town - Rural areas - Not coded	0.64%	**Y**	**Y, residency location**	**Y**
Body mass index	Continuous	Calculated according to the weight and height at the time of admission[44]	3.32%	**Y**	**N**	**N**
Smoking status	Categorical	Binary, determined according to the ICD codes	1.04%	**Y**	**N**	**Y**
Alcohol use disorder			1.05%	**Y**	**N**	**Y**
AIDS			0	**Y**	**N**	**Y**
Asthma			0	**Y**	**N**	**Y**

Cancer			0	**Y**	**N**	**Y**
Cardiovascular disorders					**N**	
Chronic heart failure					**N**	
Chronic kidney diseases			0	**Y**	**N**	**Y**
Chronic obstructive pulmonary disease			0	**Y**	**N**	**Y**
Connective tissue disease			0	**Y**	**N**	**Y**
Dementia			0	**Y**	**N**	**N**
Diabetes			0	**Y**	**N**	**N**
Dialysis			0	**Y**	**N**	**Y**
Hypertension			0	**Y**	**N**	**Y**
Immunodeficiency					**N**	
Interstitial lung disease			0	**Y**	**N**	**Y**
Leukemia			0	**Y**	**N**	**Y**
Liver failure			0	**Y**	**N**	**N**
Lymphoma			0	**Y**	**N**	**Y**
Myocardial infarction			0	**Y**	**N**	**N**
Peptic ulcer disease					**N**	
Peripheral vascular disease			0	**Y**	**N**	**Y**
Valvular dysfunction			0	**Y**	**N**	**Y**
Hemoglobin, the lowest	Continuous	The highest and lowest values from the initial day with available test results were recorded; within the first three days of admission.	2.85%	**Y**	**Y, erythrocyte count**	**N**

Hematocrit, the highest			2.84%	**Y**	**Y, erythrocyte count**	**Y**
Platelets, the highest and lowest			2.93%	**Y**	**Y, within**	**Y**
Leukocytes, the highest and lowest			2.93%	**Y**	**Y, within**	**Y**
Lymphocytes, the highest and lowest			8.71%	**Y**	**Y, within**	**N**
Neutrophils, the highest and lowest			13.44%	**Y**	**Y, within**	**N**
Monocyte, the highest and lowest			15.92%	**Y**	**Y, within**	**N**
Eosinophil, the highest and lowest			16.89%	**Y**	**Y, within**	**N**
Glucose, the highest and the lowest			5.39%	**Y**	**Y, within**	**N**
Potassium, the highest and lowest			7.52%	**Y**	**Y, within**	**N**
Sodium, the highest and lowest			7.4%	**Y**	**Y, within**	**Y**
Calcium, the highest and lowest			9.51%	**Y**	**Y, within**	**N**
Chloride, the highest and lowest			7.94%	**Y**	**Y, within**	**N**
Creatinine	Continous	The first value that is available within the first three days of admission.	5.5%	**Y**	**Y, renal function**	**Y**
Blood urea nitrogen			6.07%	**Y**	**Y, renal function**	**N**

Total bilirubin			34.06%	**Y**	**N**	**N**
ALT			32.33%	**Y**	**N**	**N**
AST			32.36%	**Y**	**N**	**N**
ALP			33.55%	**Y**	**N**	**Y**
Albumin			31.61%	**Y**	**N**	**Y**
Bicarbonate			7.04%	**Y**	**N**	**N**
Lactate			49.71%	**N**	-	-
iNR			52.78%	**N**	-	-
Prothrombin time			53.02%	**N**	-	-
Magnesium			54.88%	**N**	-	-
Direct bilurubin			60.22%	**N**	-	-
CRP			80.7%	**N**	-	-
Phosphate, the highest and lowest			87.36%	**N**	-	-
D-dimer			88.68%	**N**	-	-
Procalcitonin			89.18%	**N**	-	-
Ferritin			90.81%	**N**	-	-
LDH			92.8%	**N**	-	-
Creatinine kinase			93.58%	**N**	-	-
Fibrinogen			93.84%	**N**	-	-
Ammonia			97.24%	**N**	-	-

[*] In derivation and validation sets, combined, [**] Included in the model development studies, [***] Stayed in the final model. AIDS: acquired immunodeficiency syndrome, ALP: alkaline phosphatase, ALT: alanine transaminase, AST: aspartate aminotransferase, CRP: C-reactive protein, ICD: International Classification of Diseases, iNR: international normalized ratio, LDH: lactate dehydrogenase, N: No, VIF: Variance inflation factor, Y: Yes

Supplementary table 4. Distribution of baseline characteristics between derivation and validation datasets (before imputation).

Variables	Derivation (*n*=8043)	Validation (*n*=3451)	*P* value*	ID-validation (*n*=2441)	*P* value**
Age, years, median (IQR)	65 (49, 76)	64 (48, 76)	0.438	74 (63, 83)	**<0.001**
Sex, no. (%)			0.533		**0.009**
Female	3899 (48.5)	1695 (49.2)		110 (45.5)	
Male	4136 (51.5)	1753 (50.8)		1329 (54.5)	
Race, no. (%)			**0.011**		**<0.001**
African American	391 (4.9)	153 (4.4)		52 (2.1)	
Asian	154 (1.9)	98 (2.8)		26 (1.1)	
White	7136 (88.8)	3034 (87.9)		2299 (94.2)	
Others	355 (4.4)	166 (4.8)		64 (2.6)	
Ethnicity, no. (%)			0.627		**<0.001**
Hispanic or Latino	372 (4.6)	172 (5)		43 (1.8)	
Not Hispanic or Latino	7456 (92.8)	3194 (92.6)		2352 (96.4)	
Others, unknown, or not applicable	207 (2.6)	83 (2.4)		46 (1.9)	
Quarter of admission, no. (%)			0.783		**<0.001**
January-March	1723 (21.4)	731 (21.2)		566 (23.2)	
April-June	1750 (21.8)	777 (22.5)		803 (32.9)	
July-September	2297 (28.6)	963 (27.9)		582 (23.8)	
October-December	2273 (28.3)	980 (28.4)		490 (20.1)	
Admission location, no. (%)			0.169		**<0.001**
Arizona	1406 (17.5)	602 (17.4)		191 (7.8)	

Florida	1226 (15.2)	492 (14.3)		135 (5.5)	
MCHS	2482 (30.9)	1132 (32.8)		665 (27.2)	
Rochester	2929 (36.4)	1225 (35.5)		1450 (59.4)	
Admission source, no. (%)			0.838		**0.002**
Another hospital or care facility	1802 (22.4%)	787 (22.8)		604 (24.7)	
Outpatient or emergency department	930 (11.6%)	389 (11.3)		319 (13.1)	
Others or unknown	5311 (66%)	2275 (65.9)		1518 (62.2)	
Pre-hospital location home, no (%)	5898 (73.3)	2501 (72.5)	0.341	1617 (66.2)	**<0.001**
Transferred patient, no. (%)	1182 (20.3)	579 (22.2)	0.077	284 (13.5)	**<0.001**
Country of residence, no. (%)			0.996		0.160
United States or Canada	7990 (99.4)	3429 (99.4)		2431 (99.6)	
Others	49 (0.6)	21 (0.6)		9 (0.4)	
***African Region**	1 (2)	1 (4.8)		0	
*** Eastern Mediterranean Region**	29 (59.2)	12 (57.1)		9 (100)	
*** Region of the Americas, other than the US and Canada**	16 (32.7)	7 (33.3)		0	
*** South-East Asian Region**	2 (4.1)	1 (4.8)		0	
*** Western Pacific Region**	1 (2)	0		0	
RUCA codes, no. (%)			0.896		**0.003**
Metropolitan area	5212 (64.9)	2238 (65)		1514 (62.1)	
Micropolitan area	1100 (13.7)	475 (13.8)		410 (16.8)	
Small town	898 (11.2)	398 (11.6)		266 (10.9)	
Rural areas	783 (9.7)	317 (9.2)		240 (9.8)	
Not coded	38 (0.5)	17 (0.5)		8 (0.3)	

Body mass index, kg/m^2, median (IQR)	27.7 (23.7, 32.7)	27.9 (23.9, 32.8)	0.349	27.7 (23.7, 33)	0.554
Smoking, no. (%)			0.609		**<0.001**
Active smoker	3173 (39.5)	1379 (40)		712 (29.2)	
Never or ex-smoker	4870 (60.5)	2072 (60)		1729 (70.8)	
Alcohol use disorder, no (%)	1016 (12.8)	429 (12.6)	0.766	328 (13.4)	0.297
Comorbidities, no. (%)					
AIDS	116 (1.4)	55 (1.6)	0.539	64 (2.6)	**<0.001**
Asthma	2013 (25)	861 (25)	0.929	836 (34.2)	**<0.001**
Cancer	2939 (36.5)	1186 (34.4)	**0.026**	1229 (50.3)	**<0.001**
Cardiovascular disorders	2022 (25.1)	885 (25.6)	0.568	939 (38.5)	**<0.001**
Chronic heart failure	2119 (26.3)	889 (25.8)	0.513	1010 (41.4)	**<0.001**
Chronic kidney diseases	2440 (30.3)	1057 (30.6)	0.755	447 (18.3)	**<0.001**
Chronic obstructive pulmonary disease	1668 (20.7)	713 (20.7)	0.925	859 (35.2)	**<0.001**
Connective tissue disease	514 (6.4)	211 (6.1)	0.576	272 (11.1)	**<0.001**
Dementia	872 (10.8)	372 (10.8)	0.922	573 (23.5)	**<0.001**
Diabetes	3229 (40.2)	1377 (39.9)	0.806	1417 (58)	**<0.001**
Dialysis	448 (5.6)	188 (5.4)	0.792	1158 (47.4)	0.643
Hypertension	5314 (66.1)	2241 (65)	0.241	1974 (80.9)	**<0.001**
Immunodeficiency	773 (9.6)	329 (9.5)	0.897	265 (10.9)	0.071
Interstitial lung disease	2296 (28.6)	985 (28.6)	0.996	1208 (49.5)	**<0.001**
Leukemia	316 (3.9)	146 (4.2)	0.450	132 (5.4)	**0.002**
Liver failure	2202 (27.4)	943 (27.3)	0.954	322 (13.2)	**<0.001**
Lymphoma	405 (5)	169 (4.9)	0.755	165 (6.8)	**0.001**

Myocardial infarction	1447 (18)	626 (18.1)	0.849	631 (25.9)	**<0.001**
Peptic ulcer disease	771 (9.6)	337 (9.8)	0.765	859 (35.2)	**<0.001**
Peripheral vascular disease	2480 (30.8)	1041 (30.2)	0.476	1337 (54.8)	**<0.001**
Valvular dysfunction	2595 (32.3)	1086 (31.5)	0.402	130 (5.3)	**<0.001**
Laboratory variables at the time of admission, median (IQR)					
Hemoglobin, gr/dL	12.2 (10.2, 13.7)	12.2 (10.4, 13.8)	0.131	11.5 (9.8, 12.9)	**<0.001**
Hematocrit, %	37.5 (32.2, 41.7)	37.6 (32.7, 41.9)	0.199	36 (31.2, 39.9)	**<0.001**
Platelets, $\times 10^9$/L					
Highest	226 (169, 289)	226 (169, 286)	0.745	204 (148, 272)	**<0.001**
Lowest	222 (166, 285)	223 (166, 283)	0.783	199 (144, 266)	**<0.001**
Leukocytes, $\times 10^9$/L					
Highest	8.9 (6.5, 12.2)	8.8 (6.5, 12.1)	0.583	10.9 (7.4, 15.9)	**<0.001**
Lowest	8.7 (6.3, 11.8)	8.6 (6.3, 11.75)	0.861	10.6 (7.1, 15.1)	**<0.001**
Lymphocytes, $\times 10^9$/L					
Highest	1.18 (0.71, 1.79)	1.2 (0.73, 1.78)	0.209	0.91 (0.55, 1.38)	**<0.001**
Lowest	1.16 (0.69, 1.76)	1.18 (0.71, 1.75)	0.225	0.89 (0.53, 1.35)	**<0.001**
Neutrophils, $\times 10^9$/L					
Highest	6.29 (4.2, 9.6)	6.22 (4.18, 9.39)	0.541	8.69 (5.3, 13.12)	**<0.001**
Lowest	6.15 (4.11, 9.3)	6.1 (4.12, 9.16)	0.810	8.19 (4.81, 12.57)	**<0.001**
Monocytes, $\times 10^9$/L					
Highest	0.67 (0.46, 0.93)	0.66 (0.47, 0.92)	0.681	0.75 (0.44, 1.12)	**<0.001**
Lowest	0.65 (0.45, 0.91)	0.65 (0.45, 0.9)	0.786	0.72 (0.42, 1.09)	**<0.001**
Eosinophil, $\times 10^9$/L					

Highest	0.07 (0.01, 0.17)	0.08 (0.02, 0.17)	0.082	0.03 (0, 0.1)	**<0.001**
Lowest	0.07 (0.01, 0.16)	0.07 (0.01, 0.16)	0.065	0.02 (0, 0.09)	**<0.001**
Glucose, mg/dL					
Highest	123 (104, 162)	123 (104, 163)	0.828	137 (113, 186)	**<0.001**
Lowest	123 (104, 161)	122 (104, 161)	0.596	137 (112, 183)	**<0.001**
Lactate, mmol/L	1.6 (1.12, 2.4)	1.6 (1.1, 2.4)	0.789	1.9 (1.3, 2.9)	**<0.001**
Creatinine, mg/dL	0.96 (0.77, 1.31)	0.95 (0.76, 1.27)	0.232	1.12 (0.83, 1.59)	**<0.001**
Blood urea nitrogen, mg/dL	18 (13, 27)	18 (12.3, 26)	0.158	22 (16, 33)	**<0.001**
Potassium, mmol/L					
Highest	4.2 (3.8, 4.5)	4.1 (3.8, 4.5)	0.385	4.2 (3.8, 4.6)	**0.007**
Lowest	4.1 (3.8, 4.4)	4.1 (3.7, 4.4)	0.223	4.1 (3.7, 4.4)	**0.010**
Sodium, mmol/L					
Highest	138 (135, 140)	138 (135, 140)	0.385	137 (134, 140)	**<0.001**
Lowest	137 (134, 140)	137.7 (135, 140)	0.381	136 (133, 139)	**<0.001**
Calcium, mmol/L					
Highest	9.1 (8.7, 9.5)	9.1 (8.6, 9.5)	0.474	9 (8.5, 9.3)	**<0.001**
Lowest	9.1 (8.6, 9.5)	9.1 (8.7, 9.5)	0.652	8.9 (8.4, 9.3)	**<0.001**
Bicarbonate, mmol/L	24 (22, 26)	24 (22, 26)	0.986	23 (21, 26)	**<0.001**
Chloride, mmol/L					
Highest	101 (98, 104)	101 (98, 104)	0.471	100 (96, 103)	**<0.001**
Lowest	101 (97, 103)	101 (97, 104)	0.658	99 (95, 102)	**<0.001**
AST, U/L	28 (21, 46)	28 (20, 46)	0.375	30 (22, 51.5)	**<0.001**
ALT, U/L	23 (15, 41)	23 (15, 40)	0.742	23 (15, 41)	0.998

ALP, U/L	90 (69, 128)	89 (68, 122)	0.170	96 (74, 144)	**<0.001**
Total bilurubin, mg/dL	0.5 (0.3, 0.9)	0.5 (0.3, 0.9)	0.948	0.6 (0.4, 1)	**<0.001**
Albumin, g/dL	3.7 (3.3, 4.1)	3.8 (3.3, 4.1)	0.183	3.5 (3.0, 3.9)	**<0.001**

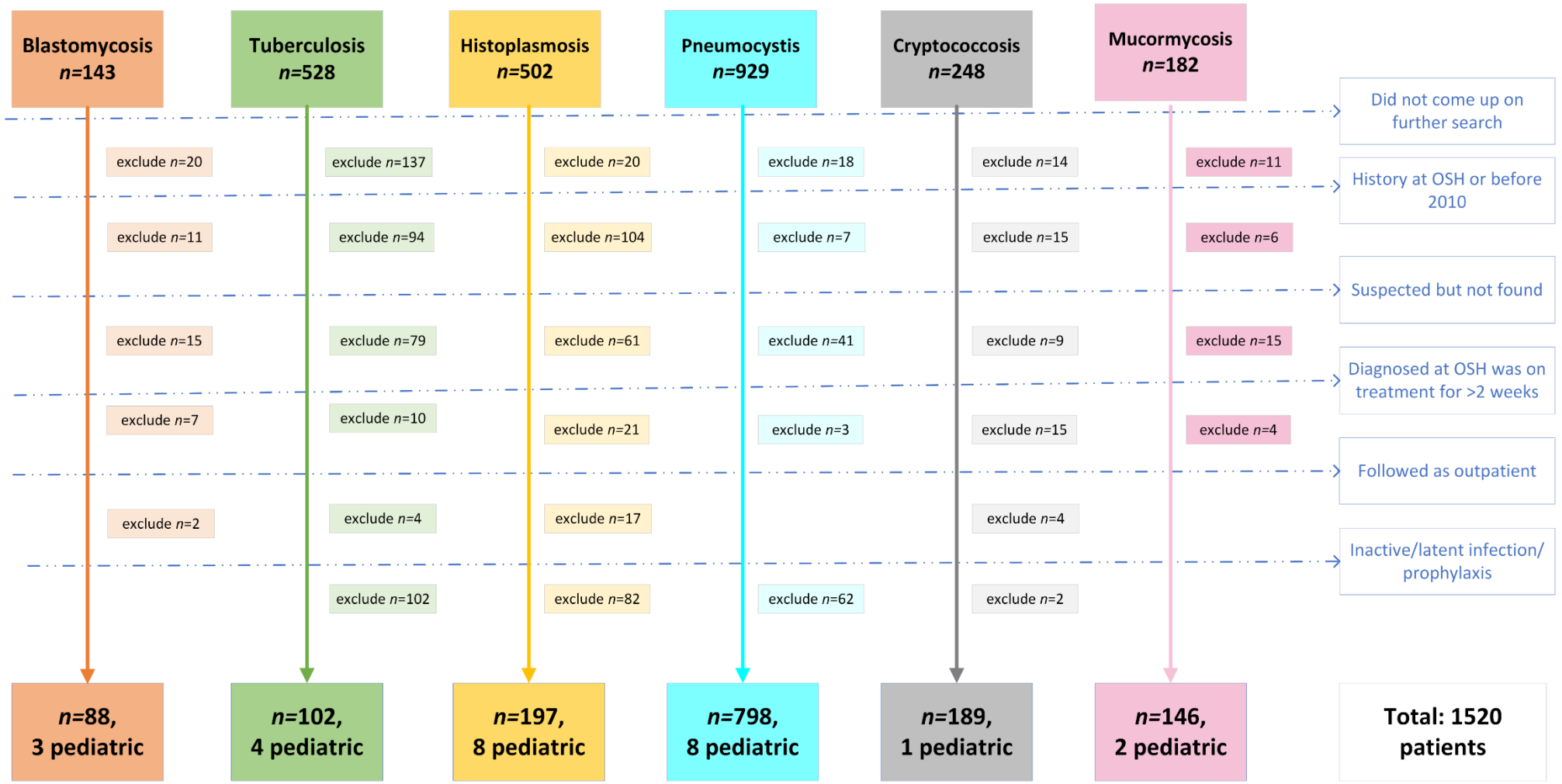
*Cases vs. controls, **Cases vs. ID-controls, ***Among those who reside outside of the United States or Canada, ****Outside of the United States or unknown.

AIDS: acquired immunodeficiency syndrome, ALP: alkaline phosphatase, ALT: alanine transaminase, AST: aspartate aminotransferase, g/dL: grams per deciliter, ID-validation: validation cases and the control group that consisted of patients with community-acquired infectious diseases other than unusual infections, IQR: interquartile range, MCHS: Mayo Clinic Health System, mg/dL: milligrams per deciliter, mmol/L: millimoles per liter, RUCA: rural-urban commuting area, U/L: units per liter.

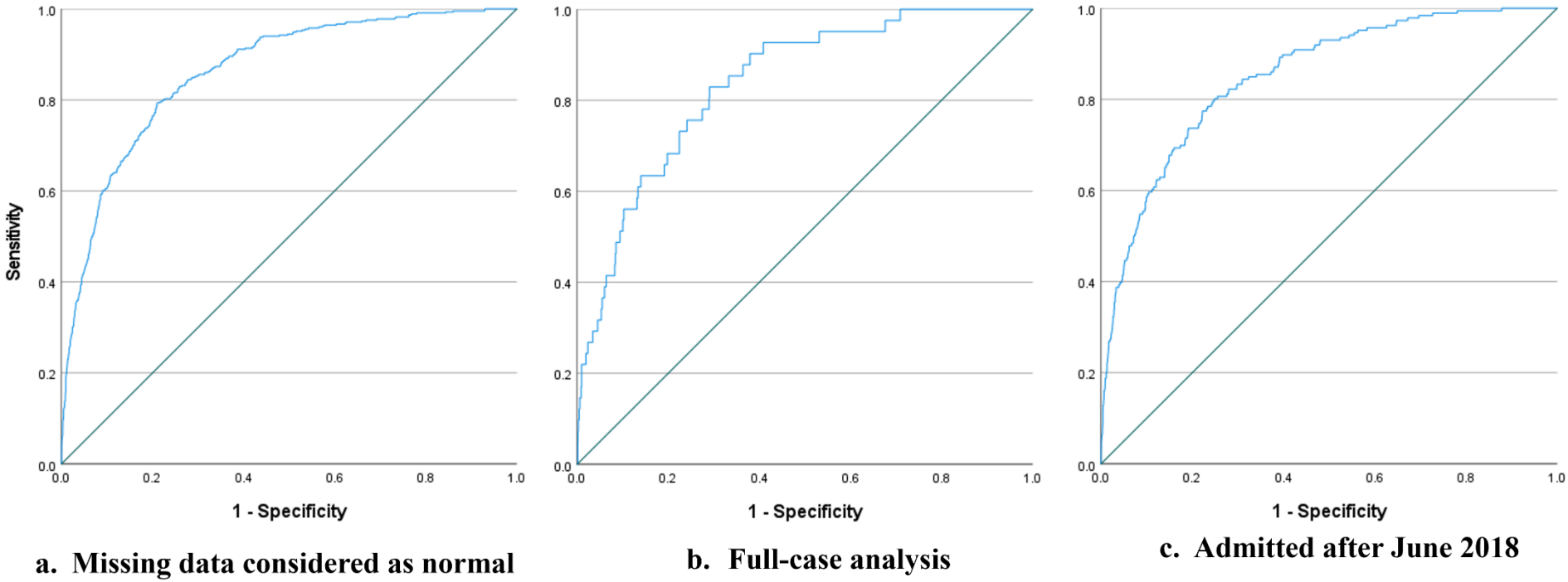
Supplementary table 5. Sensitivity analysis for different cutoff points

		Cutoff points		
		0.13	**0.10**	**0.07**
Derivation dataset	**Sensitivity (95% CI)**	0.80 (0.76-0.82)	0.85 (0.82-0.87)	0.9 (0.88-0.91)
	Specificity (95% CI)	0.80 (0.79-0.81)	0.75 (0.74-0.76)	0.66 (0.64-0.67)
	Likelihood ratio (+) (95% CI)	4.09 (3.87-4.33)	3.4 (3.24-3.57)	2.6 (2.5-2.7)
	Likelihood ratio (-) (95% CI)	0.25 (0.22-0.28)	0.21 (0.18-0.24)	0.16 (0.13-0.19)
	Positive predictive value (95% CI)	0.009 (0.008-0.009)	0.007 (0.007-0.008)	0.005 (0.005-0.006)
	Negative predictive value (95% CI)	0.999 (0.999-0.999)	0.999 (0.999-0.999)	0.999 (0.999-0.999)
	Accuracy (95% CI)	0.80 (0.80-0.81)	0.75 (0.74-0.76)	0.66 (0.65-0.67)
Validaton dataset	**Sensitivity (95% CI)**	0.69 (0.64-0.73)	0.75 (0.7-0.78)	0.87 (0.83-0.9)
	Specificity (95% CI)	0.88 (0.87-0.89)	0.85 (0.83-0.86)	0.7 (0.69-0.72)
	Likelihood ratio (+) (95% CI)	5.68 (5.06-6.37)	4.81 (4.35-5.31)	2.91 (2.72-3.11)
	Likelihood ratio (-) (95% CI)	0.36 (0.31-0.41)	0.3 (0.26-0.35)	0.19 (0.15-0.24)
	Positive predictive value (95% CI)	0.012 (0.011-0.013)	0.01 (0.009-0.011)	0.006 (0.006-0.007)
	Negative predictive value (95% CI)	0.999 (0.999-0.999)	0.999 (0.999-0.999)	0.999 (0.999-0.999)
	Accuracy (95% CI)	0.88 (0.87-0.89)	0.84 (0.83-0.86)	0.70 (0.69-0.72)
ID-validation dataset	**Sensitivity (95% CI)**	0.77 (0.72-0.80)	0.89 (0.85-0.91)	0.99 (0.98-1)
	Specificity (95% CI)	0.76 (0.74-0.78)	0.6 (0.58-0.62)	0.19 (0.17-0.21)
	Likelihood ratio (+) (95% CI)	3.18 (2.9-3.5)	2.22 (2.09-2.37)	1.23 (1.2-1.25)

	Likelihood ratio (-) (95% CI)	0.31 (0.26-0.37)	0.19 (0.14-0.24)	0.05 (0.02-0.12)
	Positive predictive value (95% CI)	0.007 (0.006-0.007)	0.005 (0.004-0.005)	0.003 (0.003-0.003)
	Negative predictive value (95% CI)	0.999 (0.999-0.999)	0.999 (0.999-0.999)	0.999 (0.999-1.00)
	Accuracy (95% CI)	0.76 (0.74-0.78)	0.60 (0.58-0.62)	0.19 (018-0.21)



Supplementary figure 1. Flowchart for the identification of study cases. OSH: Outside hospital.



Supplementary figure 2. Receiver operating characteristic of sensitivity analysis for the model for detection of patients with unusual fungal infections and tuberculosis. (A) Model performance in the validation cohort when missing data considered as normal; the area under the receiver operating characteristic curve (AUC) was 0.86 (95% CI: 0.85–0.88). (B) Model performance in the validation cohort with full-case analysis; the AUC was 0.84 (95% CI: 0.78–0.89). (C) Model performance among those admitted after June 2018; the AUC was 0.85 (95% CI: 0.83–0.88).