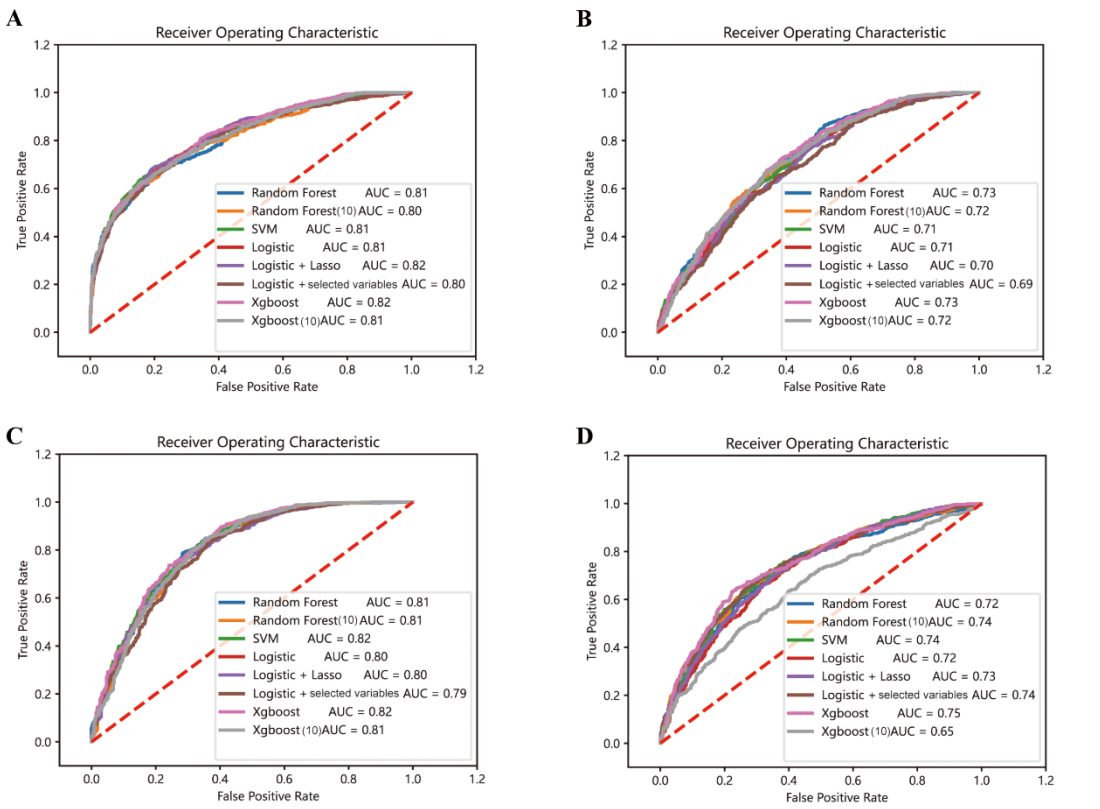
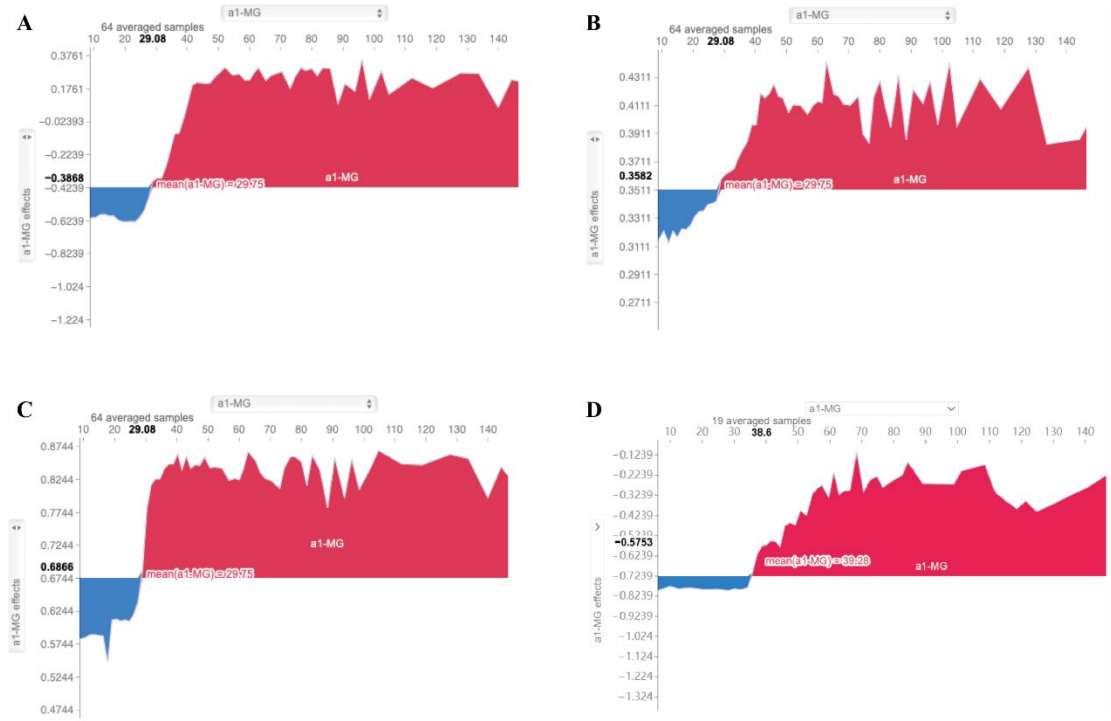


Supplemental Fig. 1 ROC curve of training and validation set for predicting DN (A), DR (B), DPN (C) and LVH (D). XGBoost: Extreme gradient boosting; SVM: Support vector machine; AUC: Area under the curve; ROC: Receiver operating characteristic.



Supplemental Fig. 2 The SHAP force plot predicting DN (A), DR (B), DPN (C) and LVH (D).

Table S1. Correlation between α1-MG levels and other clinical parameters

	Unadjusted		Age-adjusted	
	r	***P***	**r**	***P***
Age	0.124	<0.001		
SBP	0.163	<0.001	0.15	<0.001
DBP	0.069	<0.001	0.085	<0.001
BMI	0.068	<0.001	0.072	<0.001
TC	0.066	<0.001	0.076	<0.001
LDL-C	0.061	<0.001	0.069	<0.001
TG	0.112	<0.001	0.131	<0.001
HDL-C	-0.075	<0.001	-0.082	<0.001
HbA1c	0.081	<0.001	0.078	<0.001
FBG	0.05	<0.001	0.046	0.001
ALT	-0.096	<0.001	-0.076	<0.001
AST	-0.08	<0.001	-0.073	<0.001
UA	0.322	<0.001	0.345	<0.001
eGFR	-0.489	<0.001	-0.418	<0.001
CRP	0.117	<0.001	0.114	<0.001
WBC	0.178	<0.001	0.188	<0.001
LVEF	-0.068	<0.001	-0.065	<0.001

Data are presented as Pearson's correlation coefficients (r) and corresponding p-values. Both unadjusted and age-adjusted correlations are shown. SBP: Systolic blood pressure; DBP: Diastolic blood pressure; BMI: Body mass index; TC: Total cholesterol; LDL-C: Low-density lipoprotein cholesterol; TG: Triglycerides; HDL-C: High-density lipoprotein cholesterol; HbA1c: Glycated haemoglobin; FBG: Fasting blood glucose; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; UA: Uric acid; eGFR: Estimated glomerular filtration rate; CRP: C-reactive protein; WBC: White blood cell; LVEF: Left ventricular ejection fraction.

Table S2. Comparison of the performance between the models for predicting DN, DR, DPN and LVH

Outcome	Method	CI_lower	CI_upper	AUC	F1	Accuracy	Precision	Sensitivity	Specification
Diabetic nephropathy (DN)	Logistic regression training set	0.685790458	0.734272728	0.800283849	0.804702809	0.745964316	0.86557377	0.751830757	0.732525629
	Logistic regression validation set	0.700530491	0.754680197	0.81284205	0.808226221	0.753632761	0.860898138	0.761627907	0.736514523
	Logistic-Lasso regression training set	0.690851339	0.743723699	0.801579622	0.804878049	0.746247522	0.86557377	0.752136752	0.732774674
	Logistic-Lasso regression validation set	0.701694652	0.749769147	0.815141978	0.805584281	0.751651255	0.853231106	0.762977473	0.728194726
	Logistic-selected variables training set	0.680976716	0.739078925	0.797619159	0.805736636	0.746813934	0.868384075	0.751520065	0.735902256
	Logistic-selected variables validation set	0.689915951	0.750647332	0.804032345	0.800827729	0.745706737	0.847754655	0.758823529	0.718623482
	Random Forest training set	0.7569757	0.803547081	0.909285144	0.857512395	0.812800906	0.931615925	0.794329073	0.857838364
	Random Forest validation set	0.689598409	0.741341588	0.806942792	0.807535642	0.750330251	0.86856517	0.754519505	0.740820734
	Random Forest top 10 variables training set	0.754451704	0.80202343	0.899125638	0.852437703	0.807136788	0.921311475	0.793145161	0.840152236
	Random Forest top 10 variables validation set	0.692440732	0.743554481	0.804526228	0.807770961	0.751651255	0.865279299	0.757430489	0.738853503
	SVM training set	0.694967097	0.744728445	0.80169957	0.817505842	0.75672614	0.90117096	0.748055988	0.779979145

	SVM validation set	0.705485828	0.754977009	0.813574674	0.819277108	0.762219287	0.893756846	0.756255792	0.777011494
	XGBoost training set	0.706754914	0.752430788	0.8363152	0.82975137	0.77116964	0.922248244	0.754117196	0.819565217
	XGBoost validation set	0.698382613	0.746065129	**0.820244827**	0.816367265	0.756935271	0.895947426	0.749770852	0.775413712
	XGBoost top 10 variables training set	0.697838094	0.745704062	0.823637459	0.827586207	0.76635514	0.927400468	0.747169811	0.824063564
	XGBoost top 10 variables validation set	0.681772953	0.737202717	0.810386304	0.813121272	0.751651255	0.895947426	0.744313012	0.771084337
Diabetic retinopathy (DR)	Logistic regression training set	0.563647	0.622952	0.717273	0.771513	0.672897	0.8519	0.704989	0.556863
	Logistic regression validation set	0.571003	0.618964	0.705965	0.771639	0.672391	0.851626	0.705387	0.552147
	Logistic-Lasso regression training set	0.561598	0.607381	0.715941	0.78119	0.677145	0.889035	0.696679	0.583607
	Logistic-Lasso regression validation set	0.561751	0.60835	0.698389	0.777578	0.672391	0.881098	0.695827	0.563433
	Logistic-selected variables training set	0.570584	0.625736	0.707199	0.775398	0.676296	0.861948	0.704643	0.567715
	Logistic-selected variables validation set	0.559001	0.609695	0.687897	0.764273	0.661823	0.843496	0.698653	0.527607
	Random Forest training set	0.714426	0.76277	0.883718	0.858849	0.799207	0.942333	0.788954	0.834379

	Random Forest validation set	0.58171	0.635905	**0.727397**	0.777418	0.682299	0.853659	0.713679	0.5727
	Random Forest top 10 variables training set	0.694455	0.754723	0.867144	0.845224	0.782498	0.916121	0.784512	0.776224
	Random Forest top 10 variables validation set	0.591673	0.642328	0.715514	0.775281	0.682959	0.841463	0.71875	0.569061
	SVM training set	0.559084	0.605772	0.721558	0.78869	0.682526	0.913936	0.693634	0.617476
	SVM validation set	0.565039	0.607552	0.709035	0.786378	0.680978	0.903455	0.696163	0.599156
	XGBoost training set	0.61071	0.661067	0.776672	0.806677	0.717927	0.90782	0.725812	0.684132
	XGBoost validation set	0.567453	0.615753	0.726994	0.777324	0.675694	0.870935	0.701884	0.566553
	XGBoost top 10 variables training set	0.57593	0.629131	0.755074	0.795107	0.696403	0.908694	0.706762	0.644558
	XGBoost top 10 variables validation set	0.568568	0.617238	0.719407	0.78132	0.678336	0.884146	0.69992	0.579336
Diabetic peripheral neuropathy (DPN)	Logistic regression training set	0.69334	0.743646	0.804754	0.615309	0.777967	0.524686	0.743772	0.78869
	Logistic regression validation set	0.674645	0.732754	0.802059	0.593533	0.767503	0.51503	0.700272	0.789015
	Logistic-Lasso regression training set	0.685213	0.738758	0.804754	0.615309	0.777967	0.524686	0.743772	0.78869
	Logistic-Lasso regression validation set	0.670778	0.726633	0.802059	0.593533	0.767503	0.51503	0.700272	0.789015

	Logistic-selected variables training set	0.668735	0.725318	0.798011	0.585462	0.760974	0.498745	0.70868	0.777323
	Logistic-selected variables validation set	0.689483	0.740999	0.793538	0.607683	0.777411	0.523046	0.725	0.793761
	Random Forest training set	0.768284	0.809735	0.916448	0.730359	0.842538	0.630126	0.868512	0.834084
	Random Forest validation set	0.691059	0.751311	0.810624	0.622877	0.780053	0.551102	0.716146	0.80177
	Random Forest top 10 variables training set	0.755859	0.805321	0.909815	0.717255	0.834325	0.620921	0.84897	0.829507
	Random Forest top 10 variables validation set	0.69158	0.750307	0.807465	0.620047	0.784676	0.533066	0.740947	0.798268
	SVM training set	0.702622	0.757384	0.820256	0.635918	0.789861	0.542259	0.768683	0.796503
	SVM validation set	0.695314	0.748543	0.815231	0.62645	0.787318	0.541082	0.743802	0.801043
	XGBoost training set	0.736435	0.789084	0.860496	0.685657	0.821863	0.574059	0.851117	0.813211
	XGBoost validation set	0.707294	0.759821	**0.819651**	0.636792	0.796565	0.541082	0.773639	0.803433
	XGBoost top 10 variables training set	0.725263	0.779739	0.848941	0.666667	0.809969	0.561506	0.820293	0.806856
	XGBoost top 10 variables validation set	0.702092	0.758194	0.805774	0.630588	0.792602	0.537074	0.763533	0.801376

Left ventricular hypertrophy (LVH)	Logistic regression training set	0.574568	0.627927	0.724921	0.805258	0.706315	0.90426	0.725796	0.606586
	Logistic regression validation set	0.581107	0.626733	0.717768	0.790698	0.690885	0.893832	0.708901	0.606742
	Logistic-Lasso regression training set	0.590463	0.644345	0.73661	0.809758	0.715095	0.902995	0.733973	0.625407
	Logistic-Lasso regression validation set	0.589973	0.63729	0.726966	0.794397	0.699472	0.888777	0.718137	0.62069
	Logistic-selected variables training set	0.612798	0.672453	0.753226	0.808179	0.721042	0.875158	0.750724	0.614081
	Logistic-selected variables validation set	0.616393	0.679268	0.741016	0.797007	0.713342	0.861476	0.741514	0.624658
	Random Forest training set	0.70324	0.758039	0.881238	0.860646	0.797225	0.932518	0.79906	0.790576
	Random Forest validation set	0.597588	0.652321	0.724258	0.787431	0.696169	0.861476	0.725106	0.59587
	Random Forest top 10 variables training set	0.679884	0.737857	0.850614	0.83954	0.770603	0.893716	0.791558	0.704918
	Random Forest top 10 variables validation set	0.630522	0.689771	0.739395	0.798867	0.718626	0.85541	0.749336	0.628571
	SVM training set	0.643736	0.699144	0.766296	0.806236	0.728972	0.83973	0.775312	0.6054
	SVM validation set	0.635694	0.691699	0.742354	0.79307	0.715984	0.833165	0.756657	0.611765

	XGBoost training set	0.628722	0.688439	0.796887	0.827375	0.744265	0.912695	0.756643	0.691505
	XGBoost validation set	0.616406	0.669968	**0.747494**	0.801835	0.714663	0.883721	0.733837	0.643963
	XGBoost top 10 variables training set	0.540711	0.586507	0.707746	0.812567	0.700935	0.965415	0.701502	0.69403
	XGBoost top 10 variables validation set	0.531502	0.571744	0.651059	0.795474	0.677675	0.959555	0.679313	0.65812

SHAP: SHapley Additive exPlanations; CI: Confidence interval.