

Table S1 Univariate and multivariate analyses of clinical characteristics associated with overall survival in 1193 validation populations

Characteristics	Univariate analyses		Multivariate analyses	
	HR (95%CI)	P	HR (95%CI)	P
Sex (Female vs. Male)	0.53(0.44-0.63)	<0.001*	0.64(0.51-0.79)	<0.001*
Age (>59 vs. ≤59)-Year	1.66(1.35-2.04)	<0.001*	1.68(1.35-2.08)	<0.001*
Family cancer history (Yes vs. No)	0.92(0.77-1.10)	0.382	0.99(0.82-1.20)	0.931
Smoking (vs. Never)				
Ever	1.69(1.35-2.13)	<0.001*	1.17(0.89-1.55)	0.258
Current	1.55(1.25-1.91)	<0.001*	1.09(0.50-1.41)	0.497
Histology (vs. AD)				
SCC	1.51(1.25-1.82)	<0.001*	1.02(0.82-1.26)	0.868
Other [†]	1.12(0.82-1.53)	0.470	0.91(0.65-1.26)	0.554
EGFR (vs. wild type)				
Mutation	0.86(0.65-1.15)	0.320	1.20(0.88-1.62)	0.254
Unknown	0.91(0.73-1.12)	0.378	0.95(0.76-1.20)	0.688
Grade (vs. G1)				
G2	0.83(0.69-1.00)	0.046*	0.90(0.74-1.10)	0.299
G3	0.44(0.33-0.59)	<0.001*	0.54(0.40-0.74)	<0.001*
Missing	0.45(0.28-0.72)	0.001*	0.53(0.33-0.86)	0.011*
TNM stage (II vs. I)	1.85(1.54-2.23)	<0.001*	1.76(1.43-2.19)	<0.001*
Radiotherapy (With vs. Without)	1.36(0.87-2.12)	0.182	1.06(0.67-1.70)	0.797
ACT (With vs. Without)	1.11(0.90-1.36)	0.339	0.84(0.67-1.06)	0.144

[†]: Other subtype includes large cell, adenosquamous, sarcomatoid, basaloid, and unclassifiable NSCLC. *: P<0.05. AD, lung adenocarcinoma; SCC: lung squamous cell carcinoma; EGFR, epidermal growth factor receptor; ACT, adjuvant chemotherapy.

Table S2 Variants included in the polygenic hazard score construction

SNP	Locus	Function	Gene	A1 [†]	A2	Frequency [‡]	HR [§]	P*
rs17080884	5q35.3	intergenic	TRIM7; MIR4638	C	T	0.09	2.2	8.75E-06
rs7526539	1q41	intergenic	CCDC185; CAPN8	G	A	0.06	2.3	1.45E-05
rs11147267	12q24.33	intergenic	ZNF10; ZNF268	T	G	0.09	2.0	2.45E-05
rs12051246	16q21	intergenic	CFAP20; CSNK2A2	G	A	0.49	1.6	2.63E-05
rs9295494	6p22.3	intronic	CDKAL1	T	C	0.33	1.6	3.66E-05
rs9856392	3p24.2	intronic	RARB	A	G	0.40	0.6	4.59E-05
rs3803803	17q24.2	intronic	PITPNC1	T	C	0.14	1.7	8.19E-05
rs10894205	11q24.3	intergenic	ADAMTS15; MIR8052	G	C	0.26	1.6	8.41E-05
rs4751329	10q26.3	intronic	TCERG1L	C	T	0.49	1.6	1.27E-04
rs150640050	5q13.3	intronic	SV2C	G	A	0.05	2.2	1.33E-04
rs12901901	15q26.3	intergenic	LINC02348; TM2D3	T	C	0.44	1.5	1.45E-04
rs1881704	3p24.2	intronic	RARB	G	C	0.08	1.9	2.16E-04
rs7512365	1q41	intronic	ESRRG	C	G	0.50	1.5	2.19E-04
rs3857953	8q24.13	intronic	MTSS1	T	C	0.21	0.6	2.29E-04
rs2351131	6q25.2	intronic	CNKSR3	C	T	0.32	1.5	2.68E-04
rs1807209	8p23.1	intronic	ANGPT2; MCPH1	T	C	0.12	1.7	3.00E-04
rs6564755	16q23.2	intronic	LOC102724084	C	A	0.25	1.5	3.07E-04
rs35852284	12q14.3	intergenic	GRIP1; LOC102724421	G	A	0.07	2.1	3.29E-04
rs9570021	13q21.1	intergenic	LINC00374; DIAPH3	T	C	0.48	0.7	3.74E-04
rs13218513	6q22.33	intronic	SOGA3	T	C	0.05	2.1	4.34E-04
rs2969584	2q22.3	intronic	ARHGAP15	T	C	0.19	1.6	4.70E-04
rs1427897	5q23.2	intergenic	LOC101927488; GRAMD2B	C	T	0.20	0.6	4.92E-04
rs7765054	6q22.1	intergenic	HDAC2-AS2; LOC105377962	C	T	0.50	1.5	5.09E-04
rs9863866	3q13.31	intergenic	LINC00903; TUSC7	T	G	0.39	1.5	5.37E-04
rs17674822	13q32.3	intronic	LOC101927437	G	A	0.23	1.5	5.52E-04
rs766908	4q35.1	intergenic	LINC02427; LINC02365	A	G	0.21	1.6	5.95E-04
rs968763	2p25.1	intronic	LOC101929551	A	G	0.38	1.5	7.04E-04
rs71232209	12p11.21	intergenic	DDX11; FAM60A	A	G	0.06	2.1	7.22E-04
rs55928628	10q24.33	intronic	NEURL1	T	C	0.34	1.5	7.35E-04
rs12444272	16p12.3	upstream	ITPRIPL2	G	A	0.12	1.7	7.67E-04
rs7097972	10p12.2	intronic	PIP4K2A	T	C	0.43	1.5	7.75E-04
rs799169	7q11.23	intergenic	MLXIPL; VPS37D	A	G	0.20	1.5	7.85E-04
rs1294340	1q42.2	intergenic	PCNX2; MAP3K21	T	G	0.29	1.5	8.60E-04
rs838854	12q24.31	intergenic	NCOR2; SCARB1	A	G	0.40	1.4	8.68E-04
rs138451104	10q11.21	intergenic	ZNF22; RSU1P2	A	G	0.13	1.6	8.85E-04
rs62034970	16p12.3	intronic	PDILT	G	C	0.12	1.7	9.08E-04
rs9644758	8p23.1	intergenic	CTSB; DEFB136	A	C	0.48	0.7	9.49E-04

[†]: A1, Effect allele; [‡]: Frequency of Effect allele.; [§]: Hazard ratio of Effect allele from GWAS; *: P-value from GWAS.

Table S3 Baseline characteristics of patients with different PHS groups of the study cohorts

Characteristics	SYSUCC validation cohort			PLCO cohort		
	High PHS (n=463)	Low PHS (n=363)	P	High PHS (n=329)	Low PHS (n=38)	P
Sex			0.855			0.356
Male	277 (59.8%)	214 (59.0%)		178 (54.1%)	17 (44.7%)	
Female	186 (40.2%)	149 (41.0%)		151 (45.9%)	21 (55.3%)	
Age (Year)			0.728			1.000
≤59	207 (44.7%)	157 (43.3%)		25 (7.6%)	2 (5.3%)	
>59	256 (55.3%)	206 (56.7%)		304 (92.4%)	36 (94.7%)	
Family cancer history			0.015*			1.000
No	379 (81.9%)	271 (74.7%)		129 (39.3%)	15 (39.5%)	
Yes	84 (18.1%)	92 (25.3%)		199 (60.7%)	23 (60.5%)	
Smoking			0.816			0.170
Never	237 (51.2%)	191 (52.6%)		38 (11.6%)	3 (7.9%)	
Ever	80 (17.3%)	65 (17.9%)		121 (36.8%)	9 (23.7%)	
Current	146 (31.5%)	107 (29.5%)		170 (51.7%)	26 (68.4%)	
Histology			0.114			0.391
AD	313 (67.6%)	233 (64.2%)		188 (57.1%)	19 (50.0%)	
SCC	128 (27.6%)	100 (27.5%)		98 (29.8%)	11 (28.9%)	
Other [†]	22 (4.8%)	30 (8.3%)		43 (13.1%)	8 (21.1%)	
EGFR [‡]			0.192			-
Wild type	184 (39.7%)	131 (36.1%)		-	-	
Mutation	120 (25.9%)	85 (23.4%)		-	-	
Unknown	159 (34.3%)	147 (40.5%)		-	-	
Grade			0.715			0.955
G1	218 (47.1%)	162 (44.6%)		127 (38.6%)	16 (42.1%)	
G2	167 (36.1%)	145 (39.9%)		119 (36.2%)	12 (31.6%)	
G3	66 (14.3%)	48 (13.2%)		56 (17.0%)	7 (18.4%)	
Unknown	12 (2.6%)	8 (2.2%)		27 (8.2%)	3 (7.9%)	
TNM Stage			0.154			0.453
I	336 (72.6%)	246 (67.8%)		274 (83.3%)	34 (89.5%)	
II	127 (27.4%)	117 (32.2%)		55 (16.7%)	4 (10.5%)	

[†]: other subtype includes large cell, adenosquamous, sarcomatoid, basaloid, and unclassifiable NSCLC. [‡]: Data on EGFR mutation status for PLCO cohort was not available. *: P < 0.05. AD, lung adenocarcinoma; SCC, lung squamous cell carcinoma; EGFR, epidermal growth factor receptor. PHS, Polygenic Hazard Score.

Table S4 Multivariate analyses of PHS and clinicopathological characteristics associated with overall survival

Characteristics	Test set		PLCO ACT set	
	HR (95%CI)	P	HR (95%CI)	P
Sex (Female vs. Male)	0.65(0.40-1.06)	0.084	0.39(0.17-0.89)	0.026*
Age (>59 vs. ≤59)-Year	1.57(1.00-2.47)	0.052	1.09(0.30-3.96)	0.892
Grade (vs. G1)				
G2	1.25(0.79-1.99)	0.343	1.40(0.60-3.26)	0.430
G3	0.62(0.14-2.70)	0.524	1.08(0.35-3.30)	0.893
TNM stage (II vs. I)	1.25(0.79-1.98)	0.349	1.71(0.81-3.61)	0.158
PHS (Low vs. High PHS)	0.20(0.12-0.35)	<0.001*	0.52(0.07-4.01)	0.534

*: P < 0.05. ACT, adjuvant chemotherapy; PHS, Polygenic Hazard Score.

Table S5 Baseline characteristics of patients in the combined validation patients after propensity score matching

Characteristics	Matched cohort			P
	All patients n=516 (%)	Surgery alone (n=258)	Surgery + ACT (n=258)	
Sex				0.520
Male	332 (64.3%)	170 (65.9%)	162 (62.8%)	
Female	184 (35.7%)	88 (34.1%)	96 (37.2%)	
Age (Year)				0.289
≤59	235 (45.5%)	111 (43.0%)	124 (48.1%)	
>59	281 (54.5%)	147 (57.0%)	134 (51.9%)	
Family cancer history				0.200
No	374 (72.5%)	194 (75.2%)	180 (69.8%)	
Yes	142 (27.5%)	64 (24.8%)	78 (30.2%)	
Smoking				0.995
Never	207 (40.1%)	103 (39.9%)	104 (40.3%)	
Ever	112 (21.7%)	56 (21.7%)	56 (21.7%)	
Current	197 (38.2%)	99 (38.4%)	98 (38.0%)	
Histology				0.531
AD	307 (59.5%)	148 (57.4%)	159 (61.6%)	
SCC	170 (32.9%)	91 (35.3%)	79 (30.6%)	
Other [†]	39 (7.6%)	19 (7.7%)	20 (7.8%)	
EGFR				0.773
Wild type	182 (35.3%)	90 (34.9%)	92 (35.7%)	
Mutation	84 (16.3%)	45 (17.4%)	39 (15.1%)	
Unknown	250 (48.4%)	123 (47.7%)	127 (49.2%)	
Grade				0.864
G1	286 (55.4%)	139 (53.9%)	147 (57.0%)	
G2	195 (37.8%)	102 (39.5%)	93 (36.0%)	
G3	33 (6.4%)	16 (6.2%)	17 (6.6%)	
TNM Stage				0.660
I	250 (48.4%)	128 (49.6%)	122 (47.3%)	
II	266 (51.6%)	130 (50.4%)	136 (52.7%)	

[†]: other subtype includes large cell, adenosquamous, sarcomatoid, basaloid, and unclassifiable NSCLC. AD, lung adenocarcinoma; SCC: lung squamous cell carcinoma; EGFR, epidermal growth factor receptor; ACT, adjuvant chemotherapy.

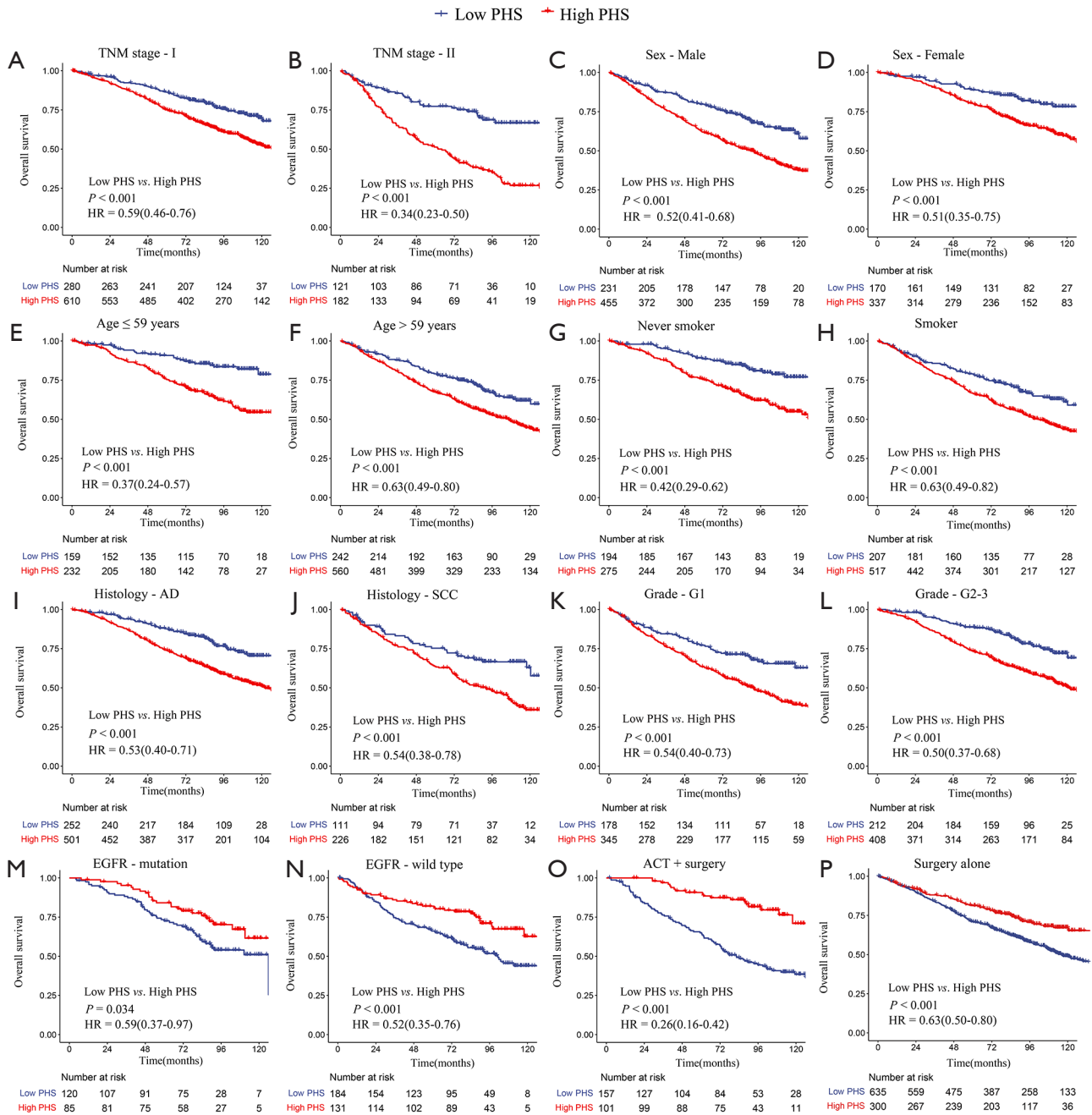


Figure S1 Kaplan-Meier analyses of overall survival by PHS in combined validation cohort stratified by clinical parameters. (A) Stage I, n=890; (B) Stage II, n=303. (C) Male, n=686; (D) Female, n=507; (E) ≤ 59 years old, n=391; (F) > 59 years old, n=802; (G) Never-smokers, n=469; (H) Former- or current-smokers, n=724; (I) Lung adenocarcinoma (AD), n=753; (J) Lung squamous cell carcinoma (SCC), n=337; (K) Grade G1 patients, n=523; (L) Grade G2-3 patients, n=620; (M) EGFR mutation, n=205; (N) EGFR wild type, n=315; (O) Patients with ACT plus surgery, n=258; (P) Patients with surgery alone, n=935. P values comparing PHS groups were calculated with the log-rank test. Hazard ratios (HRs) and 95% CIs were for low vs. high PHS in univariate COX regression analyses. PHS, Polygenic Hazard Score; EGFR, epidermal growth factor receptor; ACT, adjuvant chemotherapy.

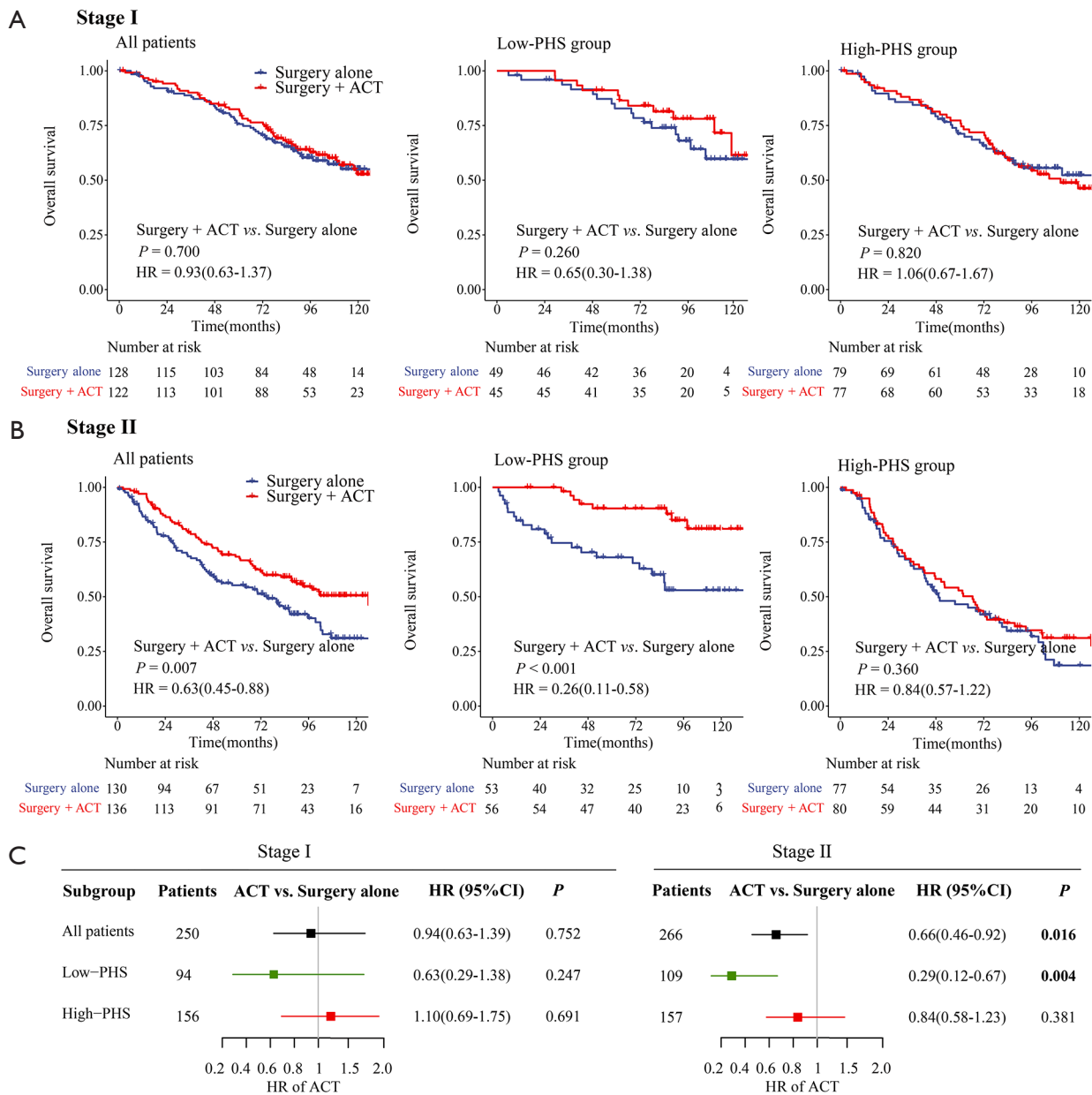


Figure S2 Kaplan-Meier and multivariate analyses of overall survival by treatment in matched cohort. (A), Kaplan-Meier plots for stage I patients for all patients (n=250, left), low-PHS group (n=94, middle), and high-PHS group (n=156, right). (B), Kaplan-Meier plots for stage II patients for all patients (n=266, left), low-PHS group (n=109, middle), and high-PHS group (n=157, right). P values were calculated with the log-rank test. Hazard ratios (HRs) and 95% CIs for surgery plus ACT vs. surgery alone were tested in univariate COX regression analyses. (C), Forest plots of HRs of ACT for stage I and stage II patients. HRs and 95% CIs were tested in multivariate Cox regression analyses adjusting for sex, age, grade, and TNM stage. ACT, adjuvant chemotherapy; Polygenic Hazard Score.