Supplementary methods for LC_MRD panel design

We developed a five-phase process to optimize genomic regions for the LC_MRD-panel. NSCLC driver genes recommended by the National Comprehensive Cancer Network (NCCN) guidelines (44), suspected NSCLC driver genes, and Chinese-specific genomic alternations in NSCLC were included in phases 1, 4, 5. In phases 2 and 3, an iterative method was used to maximize both the number of patients covered and SNVs per patient. The latter was determined by the "Recurrence Index" (RI), defined as the number of NSCLC patients harboring SNVs within a specific kilobase of exonic sequence (45). In phases 2 and 3, we analyzed non-silent SNVs identified in sequencing data of 1577 Chinese NSCLC tissue samples in the 3DMED database. Cutoffs for RI and patients per exon were chosen to enrich known or potential driver genes (45). The following steps were used to design the LC_MRD-panel.

Phase 1 (Known drivers)

Initial seed genes were chosen based upon their mutation frequency in NSCLC. COSMIC (https://cancer.sanger.ac.uk/cosmic) was used to identify known driver genes that are recurrently mutated in ≥ 9% of NSCLC. Based on the previously documented SNV pattern, certain exons from these genes were chosen. Single exons from genes with recurrent mutations that occurred at low frequency but had good evidence for being driver mutations were also included in the seed list (45-47).

Phase 2 (Max. coverage)

For each exon with a top 20% RI and a top 20% SNV frequency ranking in the 1577 Chinese NSCLC, we selected the exon with the highest RI that identified at least 1 additional patient compared to the prior phase. Among exons with equally high RI, we included the exon with the least overlap among patients already captured by the selector. The process was repeated until no more exons satisfied the aforesaid criteria.

Phase 3 (Max. median)

For each remaining exon with a top 25% RI and a top 25% SNV frequency ranking in the 1577 Chinese NSCLC, we selected the exon that would result in the biggest rise in patients with at least two SNV. For exons that are equally good, the exon with the highest RI was chosen. Repeat was performed until no other exons met the above criteria.

Phase 4 (NCCN recommended NSCLC driver genes)

We included all exons from additional genes recommended by NCCN guidelines (44).

Phase 5 (Chinese-specific mutation)

We also included genomic alternations that identified in > 0.5% of internal Chinese NSCLC tissue sample, or mutations that were previously reported (Asia WES data) for more than five times.

https://dx.doi.org/10.21037/tlcr-22-106

Table S1 Gene list of lung cancer (LC)-MRD panel

ABCB11	*CTR9*	*HOXA11*	*NOTCH1*	*RFC3*
ABL1	*CUL1*	*HRAS*	*NOTCH3*	*RGS7*
ACVR1B	*CUX1*	*IDH1*	*NRAS*	*ROS1*
ACVR2A	*DDR2*	*IGF1R*	*NRG3*	*RUNX1T1*
AFF3	*DDX10*	*IKZF1*	*NSD2*	*RXRA*
AKT1	*DDX3X*	*IL7R*	*NTRK1*	*SETBP1*
AKT2	*EGFR*	*IRS2*	*NTRK2*	*SETD2*
AKT3	*ELF3*	*JAK1*	*NTRK3*	*SFPQ*
ALK	*EP300*	*JAK3*	*PALB2*	*SH2B3*
ANK1	*EPHA2*	*KDM5A*	*PARP1*	*SLC25A13*
APEX1	*ERBB2*	*KDR*	*PAX3*	*SLC34A2*
AR	*ERBB4*	*KEAP1*	*PBRM1*	*SLIT2*
ARID1A	*ERCC3*	*KMT2A*	*PER3*	*SLX4*
ARNT	*ETV6*	*KMT2C*	*PHOX2B*	*SMAD3*
ATM	*EZR*	*KMT2D*	*PIK3CA*	*SMAD4*
ATP2B3	*FAM135B*	*KRAS*	*PIK3CB*	*SMARCA4*
ATR	*FANCA*	*LATS2*	*PIK3CD*	*SMO*
BCL11B	*FANCM*	*LIG1*	*PIK3R1*	*SOX9*
BRAF	*FAT4*	*LRP1B*	*PIK3R2*	*SPTA1*
BRCA2	*FES*	*LZTR1*	*PIK3R3*	*STAT3*
CACNA1D	*FGF3*	*MAP2K1*	*POLD1*	*STK11*
CARS	*FGFR1*	*MAP2K4*	*POLD3*	*TDG*
CBFA2T3	*FGFR4*	*MBD4*	*POT1*	*TDP1*
CBL	*FH*	*MDM2*	*PPP2R1A*	*TERT*
CCND1	*FLT1*	*MED12*	*PREX2*	*TET2*
CCNH	*FLT4*	*MET*	*PTCH1*	*TFE3*
CCNO	*FOXP1*	*MGMT*	*PTEN*	*TGFBR1*
CDH10	*FUS*	*MMS19*	*PTPN11*	*TGFBR2*
CDH11	*GATA3*	*MSH3*	*PTPN13*	*THBS2*
CDK8	*GLI1*	*MSH4*	*PTPRD*	*TIMELESS*
CDKN2A	*GLI3*	*MSH6*	*PTPRT*	*TP53*
CHD2	*GNAQ*	*MYC*	*RAD50*	*U2AF1*
CHD4	*GNAS*	*MYCN*	*RAP1GDS1*	*UGT1A1*
CLTCL1	*GRIN2A*	*MYD88*	*RASA1*	*USP8*
CNBP	*HDAC2*	*NCOR2*	*RB1*	*XAB2*
CREB3L2	*HGF*	*NF1*	*RBM10*	*XPO1*
CREBBP	*HIF1A*	*NFE2L2*	*RECQL*	*XRCC6*
CTNNB1	*HLTF*	*NFKBIA*	*RET*	*ZBTB16*
CTNND2	*HNF1A*	*NKX2-1*	*REV1*	

https://dx.doi.org/10.21037/tlcr-22-106

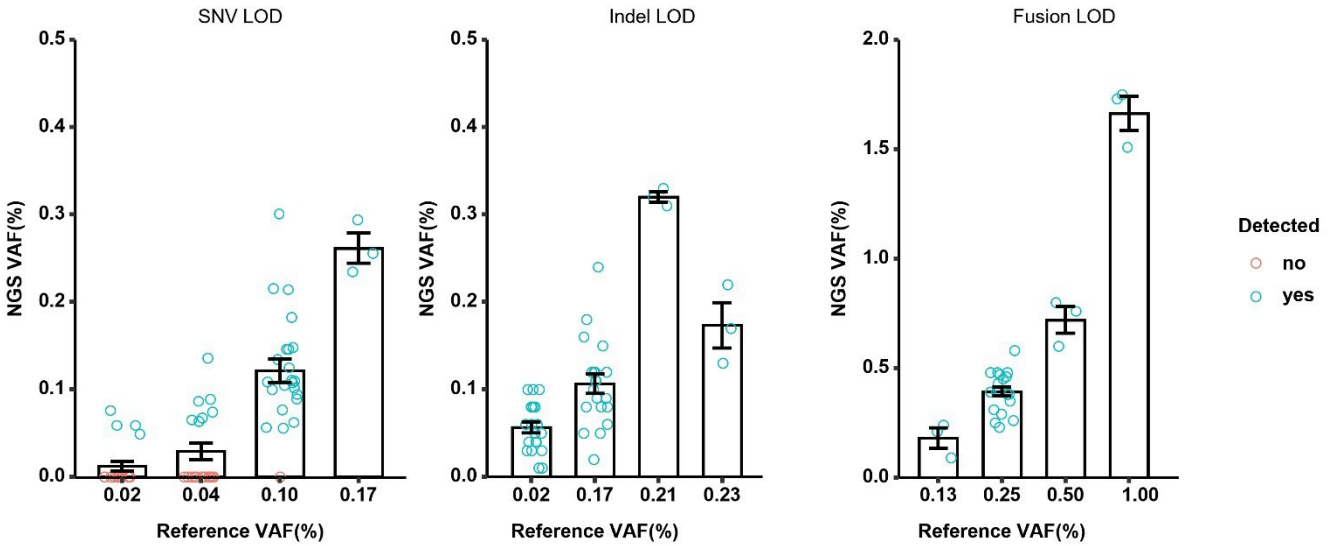

Figure S1 Limit of detection (LOD) determination for the LC-MRD assay. ctDNA reference standards of known mutations (EGFR p.L858R, EGFR p.E746_A750del, KRAS p.g12D and PIK3CA p.E545K) were diluted with wild-type ctDNA reference standard to create variant allele frequency (VAF) titration series, including 0.5%, 0.3%, 0.2%, 0.1%, and 0.05%. Genomic DNA of the tumor cell line NCI-H2228 with known ALK-EML4 rearrangement were diluted in cell line LS180 with no ALK-EML4 to generate VAF titration series of 1%, 0.5%, 0.25%, 0.125%, 0.05%. All these DNA pools' VAF were detected by digital droplet PCR (ddPCR) as a golden standard. The limit of the detection was defined as the lowest VAF at which 95% of replicates are reliably detected for the variant type. LOD of SNV/Indel and fusion were finally determined to be 0.1% and 0.25%, respectively, at 30 ng input. Values are presented as means ± standard error of mean (SEM). LOD, limit of detection; SNV, single nucleotide variant, Indel, insertions and deletions; NGS, next-generation sequencing; VAF, variant allele frequency.

https://dx.doi.org/10.21037/tlcr-22-106

Table S2 Technical validation of LC-MRD panel

Gene & amino acid sequence change	Reference VAF	Repeatability	Repeatability	Reproducibility
		batch 1	batch 2	
PIK3CA_p.E545K	2.00%	100%	100%	100%
	1.00%	100%	100%	100%
	0.50%	100%	100%	100%
	0.20%	100%	100%	100%
EGFR_19DEL	2.00%	100%	100%	100%
	1.00%	100%	100%	100%
	0.50%	100%	100%	100%
	0.20%	100%	100%	100%
KRAS_p.G12D	2.00%	100%	100%	100%
	1.00%	100%	100%	100%
	0.50%	100%	100%	100%
	0.20%	100%	100%	100%
EGFR_p.L858R	2.00%	100%	100%	100%
	1.00%	100%	100%	100%
	0.50%	100%	100%	100%
	0.20%	100%	100%	100%
ALK-EML4 fusion	2.00%	100%	100%	100%
	1.00%	100%	100%	100%
	0.50%	100%	100%	100%
	0.25%	100%	100%	100%

VAF, variant allele frequency. 19 DEL, deletion in exon 19.

https://dx.doi.org/10.21037/tlcr-22-106

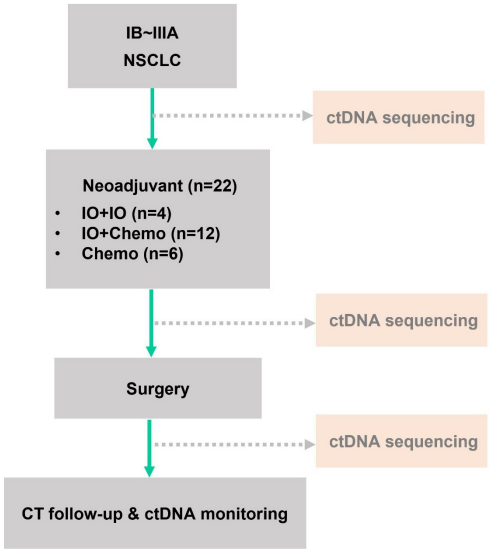

Figure S2 Schematic study design assessing the utility of ctDNA monitoring in early-stage resectable NSCLC. NSCLC, non-small cell lung cancer; IO, immunotherapy; Chemo, chemotherapy; CT, computed tomography; ctDNA, circulating tumor DNA.

https://dx.doi.org/10.21037/tlcr-22-106

Table S3 Somatic variants detected from patients' blood sample at baseline

Patient ID	Plasma used for library (mL)	Mean depth sequenced	Somatic variants (protein position)	Occurring region	Exonic function type	Allele fraction
P2	11	9611.53	EGFR_p.R831C	exonic	nonsynonymous SNV	0.001438
P3	6	10080.7	NFE2L2_p.D21H	exonic	nonsynonymous SNV	0.001074
P3	6	10080.7	ATR_p.W2379L	exonic	nonsynonymous SNV	0.00125
P3	6	10080.7	CTNNB1_p.L385I	exonic	nonsynonymous SNV	0.000963
P6	6.5	7905.32	TP53_p.E294*	exonic	stopgain	0.02344
P6	6.5	7905.32	TP53_p.C242F	exonic	nonsynonymous SNV	0.027554
P6	6.5	7905.32	CUX1_p.D1385N	exonic	nonsynonymous SNV	0.002152
P6	6.5	7905.32	EGFR_p.R831C	exonic	nonsynonymous SNV	0.001079
P6	6.5	7905.32	CDKN2A_p.D84N	exonic	nonsynonymous SNV	0.029271
P7	2.5	7662.22	EGFR_p.W817R	exonic	nonsynonymous SNV	0.003755
P7	2.5	7662.22	EGFR_p.L858R	exonic	nonsynonymous SNV	0.000631
P8	7	4921.5	HOXA11_p.G149D	exonic	nonsynonymous SNV	0.00149
P8	7	4921.5	EGFR_p.L858R	exonic	nonsynonymous SNV	0.000885
P9	3	9029.69	TP53_splicing	splicing	NA	0.001179
P9	3	9029.69	RBM10_p.Q440H	exonic	nonsynonymous SNV	0.002657
P12	5	11787.8	TP53_p.R249M	exonic	nonsynonymous SNV	0.022144
P12	5	11787.8	FAM135B_p.T1263T	exonic	synonymous SNV	0.013483
P12	5	11787.8	RUNX1T1_p.P235T	exonic	nonsynonymous SNV	0.022403
P12	5	11787.8	CDKN2A_p.R80*	exonic	stopgain	0.019832
P12	5	11787.8	RBM10_p.R744P	exonic	nonsynonymous SNV	0.012148
P13	5.5	11261.5	TP53_p.P301Qfs*44	exonic	frameshift deletion	0.007208
P13	5.5	11261.5	MBD4_p.D261G	exonic	nonsynonymous SNV	0.004162
P13	5.5	11261.5	CDKN2A_p.M54_G55insI	exonic	nonframeshift insertion	0.00247
P14	6	12032	TP53_splicing	exonic;splicing	frameshift deletion	0.114985
P14	6	12032	CTNND2_p.I420V	exonic	nonsynonymous SNV	0.00531
P14	6	12032	FAM135B_p.E544K	exonic	nonsynonymous SNV	0.067722
P14	6	12032	CDKN2A_p.V59E	exonic	nonsynonymous SNV	0.001306
P15	6.5	11338.9	KMT2D_p.P584L	exonic	nonsynonymous SNV	0.001959
P15	6.5	11338.9	TP53_p.E349*	exonic	stopgain	0.001054
P15	6.5	11338.9	TP53_p.R282W	exonic	nonsynonymous SNV	0.003833
P17	6	7549.83	TP53_splicing	splicing	NA	0.001072
P17	6	7549.83	SLC34A2_p.N495del	exonic	nonframeshift deletion	0.000926
P17	6	7549.83	EGFR_p.V441I	exonic	nonsynonymous SNV	0.002403
P18	4	8715.18	KRAS_p.G12E	exonic	nonframeshift substitution	0.000788
P18	4	8715.18	RUNX1T1_p.A461T	exonic	nonsynonymous SNV	0.001217
P19	5.5	13730.7	TP53_p.P250L	exonic	nonsynonymous SNV	0.055791
P19	5.5	13730.7	AFF3_p.R846C	exonic	nonsynonymous SNV	0.001128
P19	5.5	13730.7	CTNND2_p.A803D	exonic	nonsynonymous SNV	0.016422
P19	5.5	13730.7	FAM135B_p.Q814P	exonic	nonsynonymous SNV	0.02377
P19	5.5	13730.7	FAM135B_p.D443N	exonic	nonsynonymous SNV	0.002147
P19	5.5	13730.7	AR_p.S432F	exonic	nonsynonymous SNV	0.016431
P20	5	11069	TP53_p.G245C	exonic	nonsynonymous SNV	0.415345
P20	5	11069	NOTCH1_p.S2499Tfs*90	exonic	frameshift deletion	0.001017
P22	6	14973.9	TP53_p.R282G	exonic	nonsynonymous SNV	0.008348
P23	6.5	11881.5	KMT2D_p.E856Q	exonic	nonsynonymous SNV	0.315208
P23	6.5	11881.5	KMT2D_p.E748Q	exonic	nonsynonymous SNV	0.31836
P23	6.5	11881.5	KMT2D_p.E649K	exonic	nonsynonymous SNV	0.250229
P23	6.5	11881.5	KMT2D_p.E631K	exonic	nonsynonymous SNV	0.279321
P23	6.5	11881.5	BCL11B_p.P268L	exonic	nonsynonymous SNV	0.001996
P23	6.5	11881.5	TP53_p.E336*	exonic	stopgain	0.40132
P23	6.5	11881.5	PIK3CA_p.E545K	exonic	nonsynonymous SNV	0.009592
P23	6.5	11881.5	PIK3CA_p.H1047R	exonic	nonsynonymous SNV	0.004909
P23	6.5	11881.5	SETD2_p.F1606L	exonic	nonsynonymous SNV	0.004218
P23	6.5	11881.5	NOTCH1_p.G1753Cfs*42	exonic	frameshift deletion	0.015579
P24	5	7722.51	TP53_p.P278L	exonic	nonsynonymous SNV	0.01941
P24	5	7722.51	CDKN2A_p.W110*	exonic	stopgain	0.011224
P28	4	8380.35	TP53_p.P316Sfs*21	exonic	frameshift insertion	0.005092
P28	4	8380.35	TP53_splicing	splicing	NA	0.001103
P29	4	8552.05	TP53_p.H179R	exonic	nonsynonymous SNV	0.008401
P29	4	8552.05	PIK3CA_p.H1047L	exonic	nonsynonymous SNV	0.026549
P29	4	8552.05	CDKN2A_splicing	splicing	NA	0.006991
P30	3.5	3895.89	ARID1A_p.S593F	exonic	nonsynonymous SNV	0.001214
P30	3.5	3895.89	BCL11B_p.K875M	exonic	nonsynonymous SNV	0.000997
P30	3.5	3895.89	TP53_p.G154V	exonic	nonsynonymous SNV	0.00984
P30	3.5	3895.89	TP53_p.D61G	exonic	nonsynonymous SNV	0.001346
P30	3.5	3895.89	SLIT2_p.P1066T	exonic	nonsynonymous SNV	0.001821
P31	3.5	8195.57	GRIN2A_p.R244S	exonic	nonsynonymous SNV	0.001692
P31	3.5	8195.57	TP53_p.C238W	exonic	nonsynonymous SNV	0.001038
P31	3.5	8195.57	KEAP1_p.S233N	exonic	nonsynonymous SNV	0.001346
P31	3.5	8195.57	FAM135B_p.D555H	exonic	nonsynonymous SNV	0.001853
P37	3.5	10177.1				

SNV, single nucleotide variant; NA, not applicable.

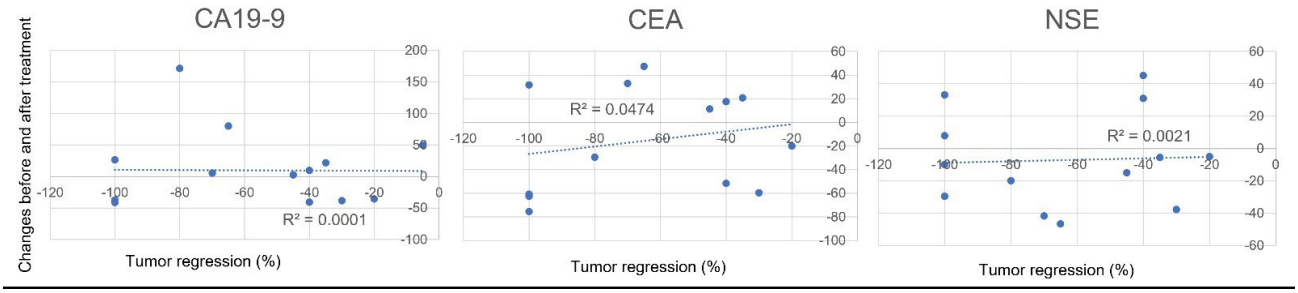

Figure S3 Correlation between pathological response and clinically used tumor biomarkers. CA19-9, serum carbohydrate antigen 19-9; CEA, carcinoembryonic antigen; NSE, neuron-specific enolase.

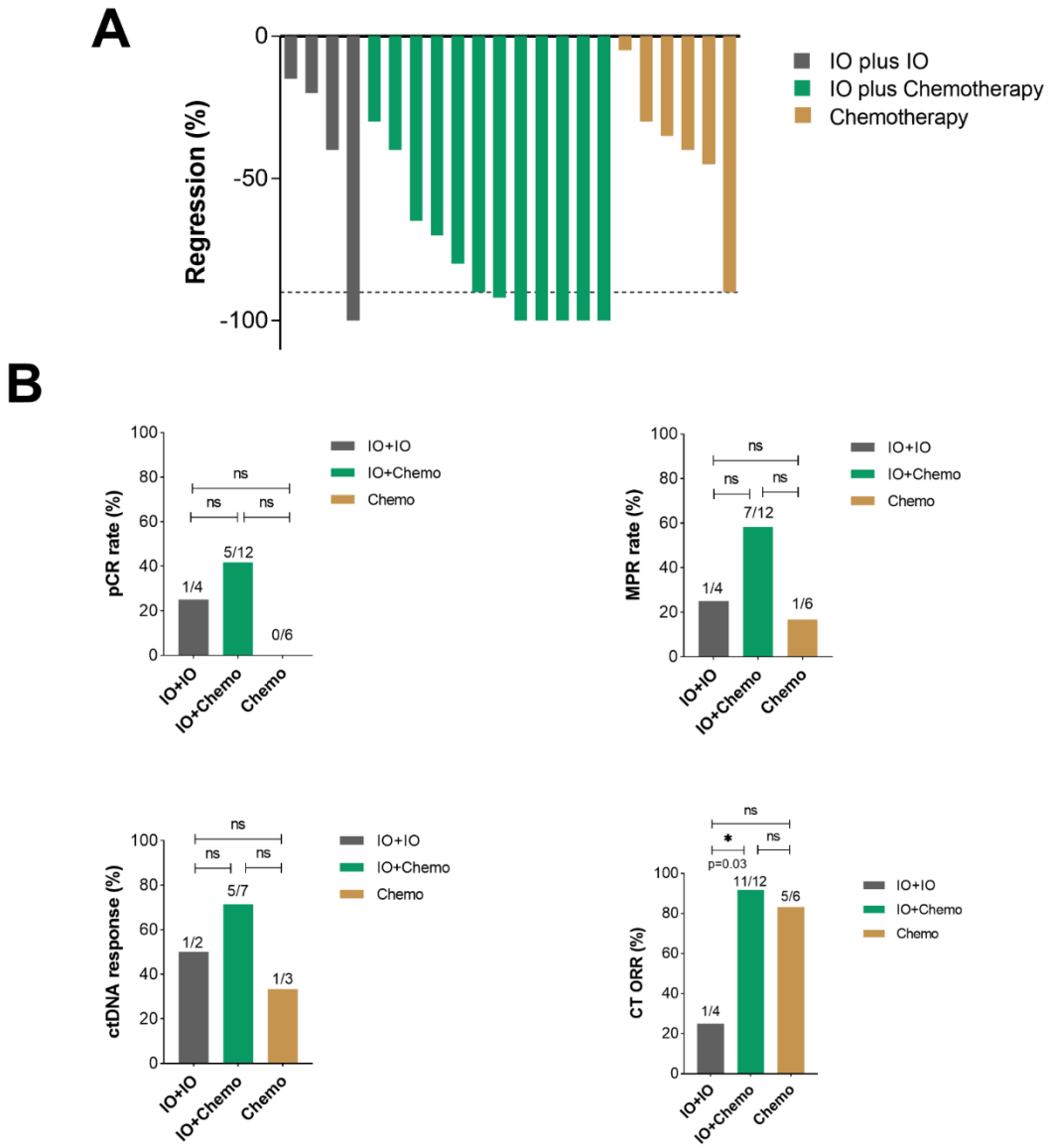

Figure S4 ctDNA change and pathological response in patients receiving different neoadjuvant regimens. (A) Tumor regression of patients treated with immunotherapy plus chemotherapy (IO+Chemo), IO+IO, and Chemo alone. (B) The pathological complete response (pCR), major pathological response (MPR), ctDNA response rates, and computed tomography (CT) response (objective response rate [ORR]) in each treatment group. ctDNA response was defined as relative delta mean variant allele fraction (RΔmean VAF) to depict dynamic changes of ctDNA upon neoadjuvant therapy. To aim for a maximum sum of specificity and sensitivity for predicting pathologic response, the threshold for defining responders was determined to 0.98. ORR was defined as the frequency of patients who have had achieved complete response (CR) or partial response (PR) at two consecutive CT assessment at least 4 weeks apart. *, P<0.05. IO, immunotherapy; Chemo, chemotherapy; IO+IO, dual immunotherapy; ctDNA, circulating tumor DNA; pCR, pathological response; MPR, major pathological response; CT, computed tomography; ORR, objective response rate.

https://dx.doi.org/10.21037/tlcr-22-106

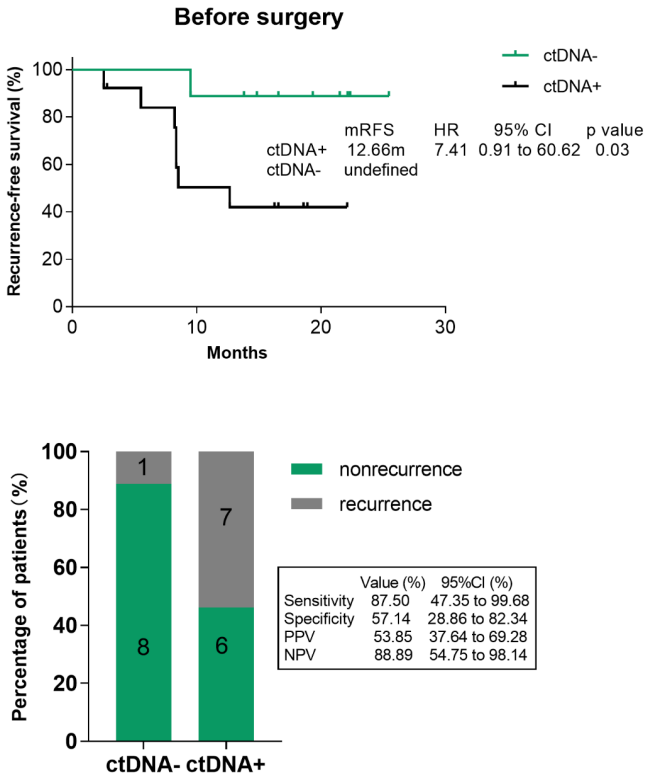

Figure S5 Absence of ctDNA after neoadjuvant therapy correlates with better recurrence-free survival (RFS). Kaplan-Meier survival curves depicting RFS of patients having had ctDNA detection after neoadjuvant therapy (within one week before surgery). HR, hazard ratio; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value. ctDNA-, without detectable mutation; ctDNA+, with detectable mutation.

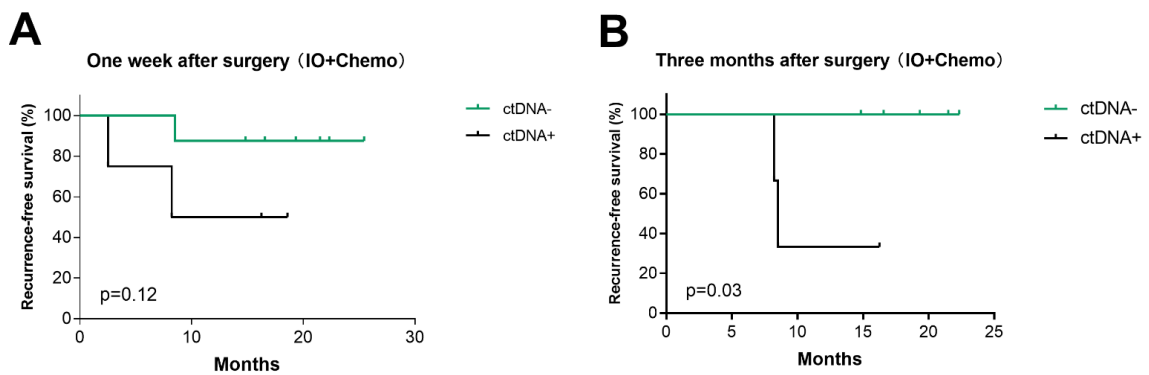

Figure S6 Correlation between postoperative ctDNA and recurrence-free survival in patients treated with immuno-chemotherapy. Kaplan-Meier survival curves depicting RFS of patients having had ctDNA detection at one week (N=22) (A) and three months (N=16) (B) after surgery. Patients who had received neoadjuvant nivolumab plus chemotherapy were analyzed. IO, immunotherapy; Chemo, chemotherapy; RFS, recurrence-free survival; ctDNA-, without detectable mutation; ctDNA+, with detectable mutation.

https://dx.doi.org/10.21037/tlcr-22-106

Table S4 Univariate and multivariate analyses of prognostic factors for recurrence-free survival

Factors	Univariate Cox			Multivariate Cox		
	HR	95% CI	P value	HR	95% CI	P value
Age (<60 *vs.* ≥60 years)	1.96	0.40–9.72	0.41			
Sex (male *vs.* female)	1.7	0.34–8.47	0.52			
Smoking (no *vs.* yes)	0.19	0.05–0.83	0.01	0.24	0.05–1.17	0.077
Histology (non-Sq *vs.* Sq)	0.44	0.11–1.75	0.23			
Disease stage (I & II *vs.* III)	0.35	0.08–1.46	0.15			
Neoadjuvant therapy						
IO plus IO *vs.* IO plus Chemo	0.27	0.04–1.66	0.16			
IO plus IO *vs.* Chemo	0.51	0.08-3.10	0.46			
*ctDNA (– *vs.* +)	5.37	1.27–22.67	0.02	4.59	1.04–20.33	0.045

*, ctDNA presence or absence in patient's blood sample at one week after surgery. HR, hazard ratio; 95% CI, confidence interval; IO, immunotherapy; Chemo, chemotherapy; Sq, lung squamous cell carcinomas.

References

44. NCCN Clinical Practice Guidelines in Oncology. Non-Small Cell Lung Cancer. Version 4. 2021 [database on the Internet]. National Comprehensive Cancer Network (NCCN). 2021. Available online: www.nccn.org.
45. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 2014;20:548-54.
46. Ding L, Getz G, Wheeler DA, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature 2008;455:1069-75.
47. Youn A, Simon R. Identifying cancer driver genes in tumor genome sequencing studies. Bioinformatics 2011;27:175-81.

https://dx.doi.org/10.21037/tlcr-22-106