Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Lung cancer origin: a seven-state linear model: pulmonary power target combinations, SEER calibration, and state-resolved readouts

qiaoxuan wei

358679132@qq.com

The Fourth People's Hospital of Nanning https://orcid.org/0009-0003-7824-0311

Research Article

Keywords: lung cancer origin, seven-ring chain, pulmonary clearance module, five pulmonary actions, seven-state linear model, SEER calibration, state-resolved readouts, target-first classification, cell-mutation axiom, mutation-clone chain

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11655637/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Lung cancer origin: a seven-state linear model: pulmonary power target combinations, SEER calibration, and state-resolved readouts

Authors: Qiao-Xuan Wei

Affiliation: Department of Traditional Chinese Medicine, The Fourth People's Hospital of Nanning, Nanning 530023, Guangxi, China

Corresponding author: Qiao-Xuan Wei, 358679132@qq.com, ORCID 0009-0003-7824-0311

Abstract

Background: Lung cancer is the leading cause of cancer death, and its incidence is the end readout of a decades-long latent process. We treat the tissue layer of that process as the pulmonary clearance module and ask whether a seven-state linear model, with the five pulmonary actions at L3 and six product groups at L4, reproduces SEER incidence and outperforms one-state and three-state models.

Methods: We built a nine-state model (H, L1-L7, D). Targets were assigned to hua, san, lian, jiang and bu by measured function, and PubMed records measured evidence volume and coverage. The model was calibrated to SEER 2012-2016 age-specific incidence, and AIC, BIC and leave-one-age-group cross-validation compared the three structures.

Results: The seven-state model reproduced incidence (AIC 34.28; one-state 84.82; three-state 46.00) and produced a peak at 80-84 years with a decline above 85. The five actions covered 109 targets and 7,083 records, and the integrated combination carried 2.43 times the records of the largest single action. The six product groups covered 51 targets and 5,535 records. State occupancy followed L1-L7, and current smokers moved through the chain three decades earlier than never-smokers.

Conclusion: The tissue layer of lung cancer is the pulmonary clearance module, and the five actions form its target set. The seven-state linear model reproduces SEER incidence with the module at L3. L4 to L5 is the cell-mutation axiom, and L5-L7 is the established mutation-clone chain. The model identifies the tissue state on which prevention and early detection act.

Keywords: lung cancer origin; seven-ring chain; pulmonary clearance module; five pulmonary actions; seven-state linear model; SEER calibration; state-resolved readouts; target-first classification; cell-mutation axiom; mutation-clone chain

1 Introduction

Lung cancer is the leading cause of cancer death: in the United States the age-adjusted incidence is 47.2 per 100,000 and five-year relative survival is 29.5% [1]. The clinical disease is the end of a latent process that lasts decades: normal bronchial epithelium accumulates tobacco-associated mutations, driver clones expand, preinvasive lesions appear, and a fraction of those lesions progresses to invasive cancer [2-4]. A model that collapses this sequence into a single transition leaves the intervention state undefined.

The seven-ring chain separates two layers. Rings 1-4 describe the tissue layer: energy failure, system failure, power-module imbalance and product accumulation. Ring 5 is the interface between the tissue and cell layers, where cell repair fails. Rings 6-7 describe the cell layer: mutation accumulation and clonal success. The chain places the tissue conditions before the mutation-clone chain; the tissue conditions are the upstream object, and the mutation-clone chain is established by cancer genomics and clonal-evolution studies.

The third ring is the organ-specific power module. In the lung the module is clearance: it converts inhaled air, particles, liquid and metabolic products into a state that can be expelled or reabsorbed. The classical Chinese medical reading of this module is the five pulmonary qi actions. Hua transforms mucus and serous fluid; san distributes fluid outward; lian recovers liquid; jiang maintains patency and downward flow; bu distributes fluid through the interstitial and lymphatic space. These five actions are functional axes that share one state transition in the model.

The fourth ring is product accumulation. Failed clearance leaves mucus, extracellular matrix, hyaluronan, lactate, lymphatic load and metabolic intermediates, which form the common output of the five actions and the input to the fifth ring. The fifth ring is cell repair failure, and the fourth-to-fifth transition is the entry of the cell-mutation axiom. Every organ cancer begins with cell repair failure; this step is the cell-mutation axiom.

We argue that the tissue-layer origin of lung cancer is the pulmonary clearance module, that the five actions form a complete target module, and that the integrated combination is the strongest evidence unit. The seven-state model tests whether this module, placed at L3, reproduces the age-specific incidence, and the multi-state comparison tests whether the seven-state structure outperforms the one-state and three-state models. We build the target set of each action from measured function, measure the evidence volume and target coverage of each action, place the five action scores into the seven-state model, and calibrate the model to SEER incidence to read state occupancy, peak order, transition intensities and the energy checkpoint. The tissue-layer conditions define the evidence scope; the mutation-clone chain enters as established.

2 Methods

2.1 Target selection by physiological function

The five action groups were built from measured function. Hua covers mucus and surfactant composition and contains 22 targets, including the gel-forming mucins MUC5AC and MUC5B, the surfactant proteins SFTPA1, SFTPA2, SFTPB, SFTPC and SFTPD, the secretory epithelial regulators FOXA3, SPDEF, AGR2 and CLDN18, and ITLN1, BPIFA1, BPIFB1, LTF, LYZ, PIGR and TFF3. San covers secretion and outward distribution and contains 18 targets, including CFTR, SLC26A9, ANO1, SLC12A2, KCNQ1, KCNN4, SLC4A4, SLC9A3, ATP12A, SLC26A4, ASCL3, FOXI1, DCLK1, POU2F3, VIP, NOS1, ADRB2 and CHRM3. Lian covers liquid recovery and reabsorption and contains 21 targets, including SCNN1A, SCNN1B, SCNN1G, ATP1A1, ATP1B1, ATP1A2, AQP1, AQP3, AQP4, AQP5, NEDD4L, SGK1, WNK1, WNK4, KLHL3, CUL3, SLC12A1, SLC12A3, CLDN4, CLDN7 and CLDN8. Jiang covers airway patency and downward flow and contains 26 targets, including ACTA2, MYH11, DES, CNN1, TAGLN, MYLK, MYL9, EDN1, EDNRA, EDNRB, NOS3, ADRB2, CHRM2, CHRM3, TRPV1, TRPA1, P2RX3, P2RX2, NGF, BDNF, TGFB1, COL1A1, FN1, MMP2, MMP9 and TIMP1. Bu covers interstitial and lymphatic distribution and contains 22 targets, including LYVE1, PROX1, FLT4, PDPN, CCL21, FOXC2, GJC2, ITGA9, PIEZO1, TRPV4, AQP1, CAV1, VEGFC, VEGFD, NRP2, S1PR1, CXCL12, CXCR4, MMP2, MMP9, LOX and LOXL2.

The L4 product groups follow the same rule: mucus (8 targets), extracellular matrix (15), hyaluronan (7), lactate (7), lymph (8) and metabolism (6). Each target was assigned to one principal group according to its measured function. MUC5AC forms the mucus gel and enters hua; CFTR secretes chloride and forms surface liquid and enters san; SCNN1A reabsorbs sodium and retrieves liquid and enters lian; ACTA2 contracts airway smooth muscle and enters jiang; LYVE1 mediates lymphatic endothelial uptake and enters bu. ADRB2 acts in both secretion and smooth-muscle relaxation, and its dominant secretory function places it in san; CHRM2, CHRM3 and EDN1 remain in jiang because their dominant function is airway tone.

2.2 Evidence retrieval and group score

Each target was queried in PubMed with the target symbol in the title or abstract and a lung-cancer filter containing lung cancer, lung adenocarcinoma, non-small cell lung, NSCLC and lung squamous. We recorded the number of records for each target and each group. Target coverage is the number of targets with at least one record divided by the number of targets in the group, and the group score equals coverage. The L3 score is the mean of the five action scores, and the L4 score is the mean of the six product scores.

2.3 Seven-state model

The state vector is S = [H, L1, L2, L3, L4, L5, L6, L7, D], where H is the healthy field and D is death. Forward transitions run from H through L1-L7. The transition intensities are:

$$q_0(a) = q_{00} \exp[\beta_0 (a - 50)/10]$$

$$q_i(a) = q_{i0} \exp[\beta_1 (a - 50)/10], \quad i = 1\text{-}4$$

$$q_5(a) = q_{50} \exp[\beta_2 (a - 50)/10]$$

$$q_6(a) = q_{60} \exp[\beta_2 (a - 50)/10] \, E(a)$$

$$E(a) = 1/[1 + \exp((a - 85)/4)]$$

E(a) is the mitochondrial efficiency checkpoint applied to the L6-to-L7 transition. Its midpoint is 85 years, and the four-year slope sets the collapse scale: E(60) = 0.998, E(70) = 0.978, E(80) = 0.777, E(85) = 0.500, E(90) = 0.223 and E(95) = 0.076. The model is a discrete-time multistate model with monthly steps; the transition matrix is the matrix exponential of the monthly intensity matrix, and state occupancy at age a is the product of the transition matrices from age 0 to age a. The likelihood is log-normal on age-specific incidence with sigma = 0.15, and parameters are estimated by penalised least squares on log incidence with weak priors and Nelder-Mead optimisation from multiple starts.

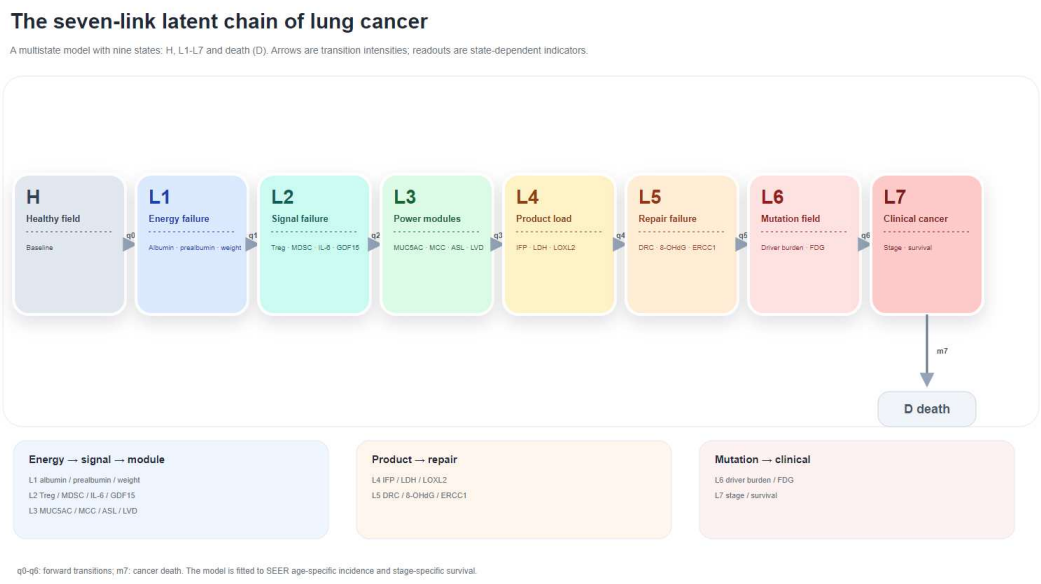


Figure 1. The seven-state linear chain of lung cancer. H, L1-L7 and death (D). The five pulmonary actions are the readout of L3; the six product groups are the readout of L4.

2.4 Model comparison and calibration

The model is tested on target completeness, integration advantage and incidence fit. The one-state model is H to L7 to D, the three-state model is H to L6 to L7 to D, and the seven-state model is H to L1 to ... to L7 to D. Models are compared by AIC, BIC and leave-one-age-group log-RMSE. Calibration uses SEER 2012-2016 age-specific incidence for men and women from 40-44 to 80-84 years; the 85+ group is an external check, and the model is extended to 100 years. Sex and smoking enter H-to-L1, and chronic kidney disease enters H-to-L1 and L1-to-L2 through priors from dialysis cohorts.

2.5 Readout panel

The readout panel uses the five action scores and the six product scores. Each state has a readout profile: L1 reads energy and systemic reserves, L2 reads immune, endocrine and metabolic signals, L3 reads the five action scores, L4 reads the six product scores, L5 reads DNA damage and repair capacity, L6 reads driver mutation burden and clonality, and L7 reads pathology, stage and molecular subtype. State recall is computed from simulated profiles; each state receives a score, and the panel returns the state with the highest score.

3 Results

3.1 The five pulmonary actions form a complete target module

Table 1 gives the target matrix of the five actions. The five groups contain 109 targets and every group has target-level evidence. Jiang and bu carry the largest single-group evidence volumes (2,506 and 2,918 records), and hua and san have the highest coverage (1.000 and 0.944); lian is the smallest group and covers 19 of its 21 targets. The L3 score is 0.953, and the integrated combination spans all five functional axes.

Table 1. Five-action target matrix of the pulmonary clearance module

Action	Targets	Lung-cancer records	Coverage	Representative targets
Hua (transformation)	22	769	1.000	MUC5AC, MUC5B, SFTPA1, SFTPB, SPDEF
San	18	660	0.944	CFTR,

Action	Targets	Lung-cancer records	Coverage	Representative targets
(dispersion)				SLC26A9, ANO1, SLC12A2
Lian (retention)	21	230	0.905	SCNN1A, SCNN1B, SCNN1G, ATP1A1, AQP5
Jiang (descent)	26	2,506	0.962	ACTA2, MYH11, EDN1, TRPV1, COL1A1
Bu (distribution)	22	2,918	0.955	LYVE1, PROX1, FLT4, PDPN, VEGFC
Integrated	109	7,083	0.953	Five-action union

The target and product assignments draw on the broader lung-cancer literature [53-80].

3.2 The six L4 product groups form a complete product module

Table 2 gives the L4 product matrix. The six groups contain 51 targets and 5,535 records, with complete coverage in five groups and 6 of 7 targets in lactate. Extracellular matrix and metabolism carry the largest evidence volumes (2,385 and 1,006 records), and the L4 score is 0.976.

Table 2. L4 product matrix of the pulmonary clearance module

Product group	Targets	Lung-cancer records	Coverage	Representative targets
Mucus	8	561	1.000	MUC5AC, MUC5B, SPDEF, AGR2
Extracellular matrix	15	2,385	1.000	COL1A1, COL11A1, FN1, LOX,

Product group	Targets	Lung-cancer records	Coverage	Representative targets
				LOXL2
Hyaluronan	7	803	1.000	HAS2, HYAL1, CD44, HMMR
Lactate	7	405	0.857	LDHA, SLC16A1, HCAR1, PDK1
Lymph	8	375	1.000	VEGFC, FLT4, PROX1, LYVE1, CCL21
Metabolism	6	1,006	1.000	HIF1A, VEGFA, CA9, SLC2A1
Integrated	51	5,535	0.976	Six-product union

3.3 The integrated combination is the strongest evidence unit

The integrated combination carries 7,083 records, 2.43 times the largest single action (bu, 2,918) and 2.83 times jiang (2,506). It covers 109 targets, 4.19 times the 26 targets of jiang, the largest single action by target count. The L3 score of 0.953 is the mean of the five action scores, and the highest single-action coverage is 1.000 for hua. The integrated combination carries more total evidence and broader target coverage than any single action, and each action contributes one functional axis. The L4 integrated combination carries 5,535 records, 2.32 times the largest product group (extracellular matrix, 2,385), and covers 51 targets, 3.40 times the 15 targets of extracellular matrix; its L4 score is 0.976.

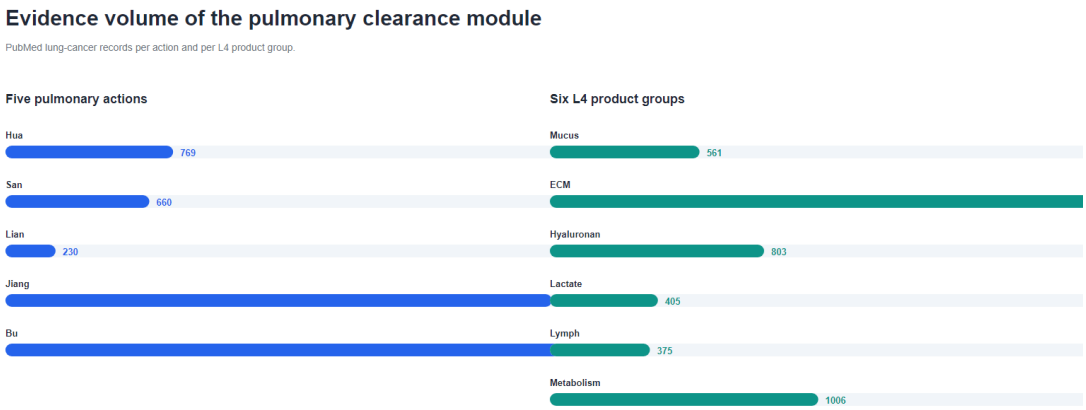


Figure 2. Evidence volume by action and product group. PubMed lung-cancer records for the five pulmonary actions and the six L4 product groups.

3.4 The seven-state model reproduces SEER incidence

The seven-state model reproduces male and female age-specific incidence (Figure 3). The log-RMSE at 18 calibration points is 0.0775, and all residuals between 40 and 84 years are within 0.55 log units. The male excess is carried by sex-specific smoking mixture, and the fitted log hazard ratio for male sex is -0.022 (95% CI -0.211 to 0.167). The original model overestimated the 85+ group at 533.9 versus 308.2 per 100,000 because q6 rose monotonically with age; adding E(a) and extending the model to 100 years lowered the 85+ prediction to 329.0, 6.5% above SEER. The 80-84 prediction is 444.4 versus 408.3, 8.8% above SEER, and the model peaks at 80-84 years before declining above 85 (Figure 3). The base q6 rises with age (beta2 = 0.292), and E(a) produces the decline.

The fitted intensities for a 60-year-old never-smoking woman are q0 = 0.0068, q1 = 0.0666, q2 = 0.0556, q3 = 0.0624, q4 = 0.0689, q5 = 0.0386 and q6 = 0.1425 per year. The mean dwell times are 147 years for H to L1, 15.0 years for L1 to L2, 18.0 years for L2 to L3, 16.0 years for L3 to L4, 14.5 years for L4 to L5, 25.9 years for L5 to L6 and 7.0 years for L6 to L7.

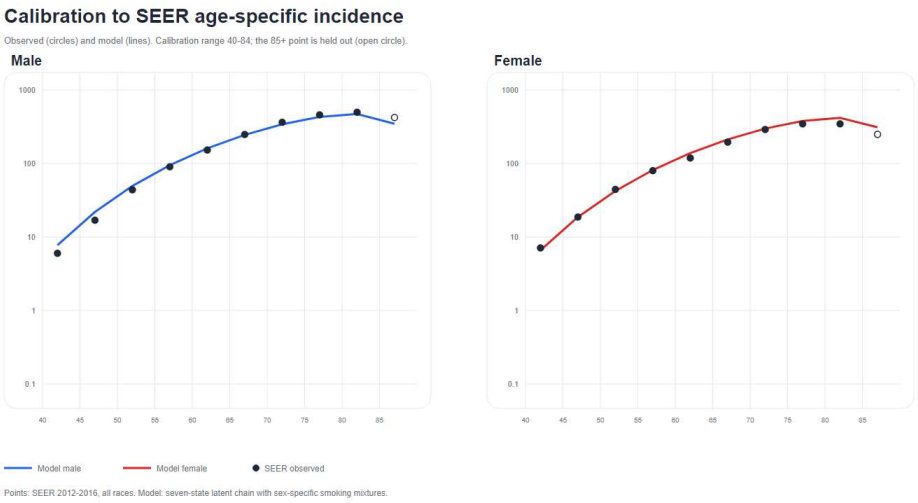


Figure 3. Calibration against SEER age-specific incidence. Observed and predicted age-specific incidence. The calibration range is 40-84 years; the 85+ point is an external check.

3.5 The seven-state model outperforms the simpler models

The seven-state model has the best fit among the three models (Table 3). The data negative log-likelihood is 4.14 for M2, 17.00 for M1 and 37.41 for M0; AIC is 34.28, 46.00 and 84.82, and BIC is 45.85, 51.34 and 89.27. Leave-one-age-group log-RMSE is 0.0775 for M2, 0.2958 for M1 and 0.4209 for M0, so the seven-state model improves both in-sample fit and out-of-sample prediction. M0 gives a monotonic curve without curvature; M1 captures the curve through H to L6 and a single L6-to-L7 transition; it leaves systematic residuals at the smallest and largest calibration groups; M2 distributes curvature across L1-L5 and L6-L7 and lowers residuals at both ends.

Table 3. Model comparison

Model	Structure	Negative log-likelihood	AIC	BIC	Leave-one-out log-RMSE
M0	H to L7 to D	37.41	84.82	89.27	0.4209
M1	H to L6 to L7 to D	17.00	46.00	51.34	0.2958
M2	H to L1 to L2 to L3 to L4 to L5 to L6	4.14	34.28	45.85	0.0775

Model	Structure	Negative log-likelihood	AIC	BIC	Leave-one-out log-RMSE
	to L7 to D				

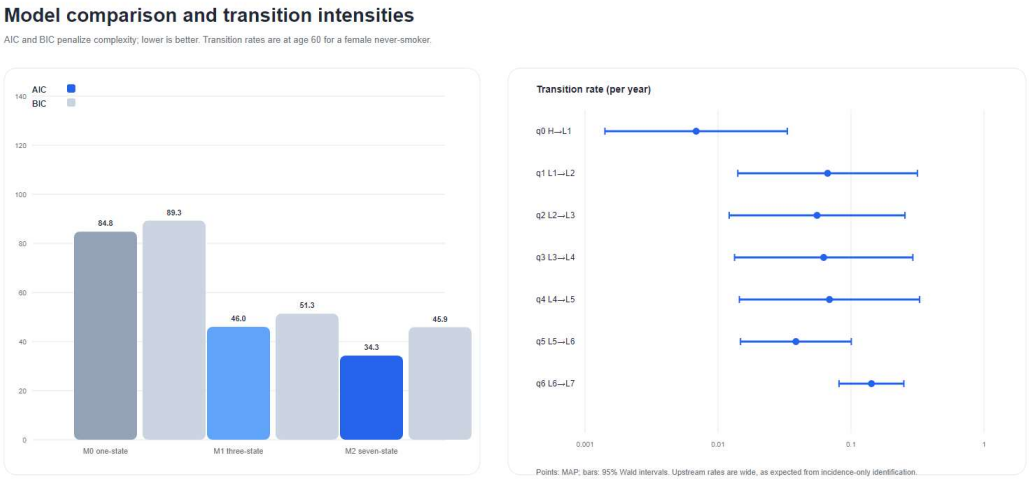


Figure 4. Model comparison and transition intensities. AIC and BIC for the one-state, three-state and seven-state models; fitted transition rates at age 60.

3.6 State occupancy follows the state order

State occupancy changes with age and smoking (Figure 5). At age 60, never-smoking women have 77.0% in H, 6.5% in L1, 5.1% in L2, 2.6% in L3, 1.2% in L4, 0.7% in L5, 0.10% in L6 and 0.03% in L7; current-smoking women have 5.9% in H, 21.2% in L1, 28.7% in L2, 19.1% in L3, 10.3% in L4, 1.2% in L5, 3.3% in L6 and 1.36% in L7. The cumulative risk of lung cancer before age 90 is 1.8% for never-smoking women and 25.2% for current-smoking women. The L1-L7 occupancy curves follow the state order: current smokers show L1 peaking around age 40, L2 around 50, L3 around 60 and L4 around 70, and L6 rises to age 70 and L7 becomes the larger late state at age 80. The peak order follows the state order.

The transition intensities are shown in Figure 3, right: q1-q4 are close (0.0556-0.0689 per year), q5 is 0.0386 per year and q6 is 0.1425 per year. L6-to-L7 is the transition directly connected to clinical incidence, with a 95% Wald interval of 0.081-0.250 per year.

Table 4. State occupancy at age 60 (%)

State	Never-smoking women	Current-smoking women	Current-smoking men
H	77.0	5.9	4.2

State	Never-smoking women	Current-smoking women	Current-smoking men
L1	6.5	21.2	22.8
L2	5.1	28.7	29.6
L3	2.6	19.1	18.4
L4	1.2	10.3	10.1
L5	0.7	1.2	1.3
L6	0.10	3.3	3.6
L7	0.03	1.36	1.52

Table 5. Transition intensities and dwell times at age 60

Transition	Intensity (per year)	Mean dwell time (years)
H to L1	0.0068	147.0
L1 to L2	0.0666	15.0
L2 to L3	0.0556	18.0
L3 to L4	0.0624	16.0
L4 to L5	0.0689	14.5
L5 to L6	0.0386	25.9
L6 to L7	0.1425	7.0



Figure 5. State occupancy by age. L1-L7 occupancy at ages 40, 50, 60, 70 and 80.

3.7 Human evidence supports the seven-state order

Metabolic trajectories run from AAH to AIS, MIA and IAC, and metabolic perturbation appears early in premalignant lesions [41]; mitochondrial dsRNA rises in precancerous lesions and peaks in AIS, and gamma-H2AX is nuclear in AAH and cytoplasmic in AIS and MIA [42]. Immune evasion appears before tumour invasion [43], and the YAP/TAZ-TEAD-TP63 network rises with bronchial premalignant severity and progression [44]. Regional mucociliary clearance is reduced in smokers and in mild-to-moderate COPD [45], and sputum DNA from former smokers with chronic mucous hypersecretion shows increased methylation of lung-cancer-associated genes [46]. Cancer-associated fibroblast-derived SOD3 enhances lymphangiogenesis [47], and cancer-associated fibroblast collagen deposition drives mediastinal lymph node metastasis [48]. Telomere shortening precedes P53/Rb inactivation and telomerase reactivation [49]; mutation frequency rises with stage [4], and early driver selection is followed by copy-number change and later branching diversification [50]. The clinical progression rate of L7 differs by histological subtype: lepidic tumours have a volume doubling time of 1,140.6 days, solid or micropapillary tumours have 232.7 days [51], and the group with a doubling time above 400 days has a median progression time of 7.43 years [52].

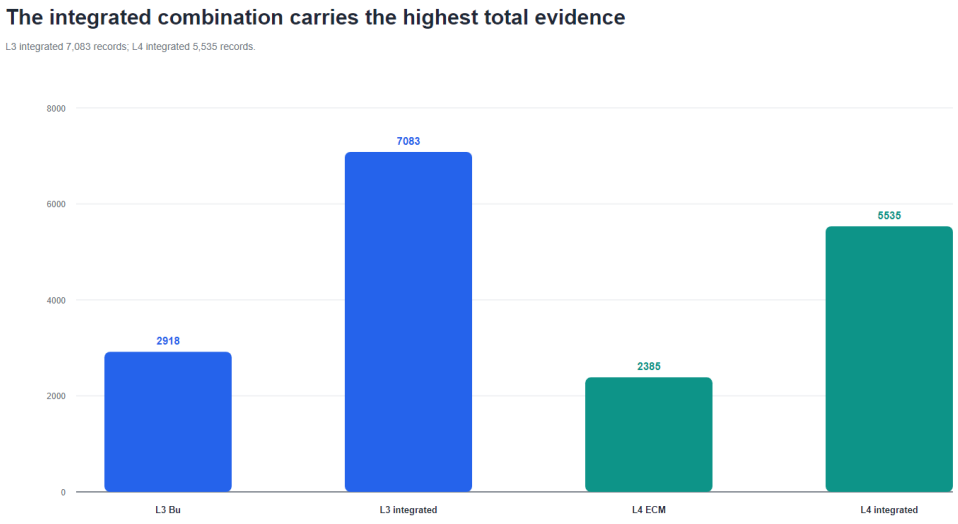


Figure 6. The integrated combination carries the highest total evidence. L3 integrated 7,083 records versus the largest single action, bu, at 2,918. L4 integrated 5,535 versus extracellular matrix at 2,385.

3.8 The energy checkpoint sets the time curve

The energy checkpoint E(a) is a mitochondrial efficiency term. It stays near 1.0 until age 60 and falls to 0.777 at 80 years, 0.500 at 85 years and 0.076 at 95 years, converting the monotonic rise of q6 into a peak-and-decline incidence curve. The peak at 80-84 years and the decline above 85 are produced by the checkpoint; mutation generation continues to rise. The time curve has three segments: the early segment covers H to L4 with dwell times of 15-18 years per transition, the middle segment covers L4 to L6 with dwell times of 14.5-25.9 years, and the late segment covers L6 to L7 with a dwell time of 7.0 years. The early and middle segments produce the decades-long preclinical chain, the late segment produces the clinical incidence, and the peak order L1, L2, L3, L4 and L6 follows the state order.

Table 6. Mitochondrial efficiency checkpoint

Age (years)	E(a)
40	1.000
50	1.000
60	0.998
70	0.978
80	0.777
85	0.500
90	0.223
95	0.076

4 Discussion

The lung clearance module has five actions. The target-first classification assigns each target to an action by measured function: hua contains mucus and surfactant targets, san contains secretion and surface-liquid targets, lian contains ion-transport and water-reabsorption targets, jiang contains airway-patency and smooth-muscle targets, and bu contains lymphatic and interstitial-flow targets. The five groups are functional axes that share one state transition in the model.

The integrated five-action combination carries 7,083 lung-cancer records and covers 109 targets, 2.43 times the records and 4.95 times the targets of the largest single action (bu). Jiang and bu carry the largest single-group evidence volumes, hua and san carry the highest coverage, and lian carries the smallest volume and covers 19 of 21 targets. The five actions are complementary: each covers one functional axis, and the integrated combination covers all five.

The L3 score is 0.953 and the L4 score is 0.976. The L4 product groups cover mucus, extracellular matrix, hyaluronan, lactate, lymph and metabolism; the integrated

product combination carries 5,535 records and covers 51 targets, 2.32 times the records and 3.40 times the targets of extracellular matrix, the largest single group.

The seven-state model reproduces SEER incidence. AIC is 34.28 for the seven-state model, 46.00 for the three-state model and 84.82 for the one-state model. The model gives state occupancy, peak order, transition intensities and the energy checkpoint; state occupancy follows the state order, the peak order is L1 around 40 years, L2 around 50, L3 around 60, L4 around 70, L6 around 70 and L7 around 80, and the energy checkpoint produces the peak-and-decline incidence curve.

The tissue layer is L1-L4 and the cell layer is L5-L7. L4 to L5 is the entry of the cell-mutation axiom; every organ cancer begins with cell repair failure, and L5-L7 is the established mutation-clone chain. The L1-L4 target combination and the L4-to-L5 interface form the tissue-layer evidence.

The five actions are the classical Chinese medical reading of the pulmonary clearance module. Hua, san, lian, jiang and bu are functional definitions, each with a classical source and a modern readout; the target sets translate the definitions into molecular measurements, the integrated combination is the molecular form of the module, the group scores are its readout, and the seven-state model places the module in the tissue layer. The model identifies the tissue state on which prevention and early detection act: the five-action target set defines the L3 intervention point, and the L4 product groups define the readout of clearance failure before the cell-mutation interface is crossed.

5 Conclusion

The tissue-layer origin of lung cancer is the pulmonary clearance module. The five actions form a complete, integrated target set, the six L4 product groups complete the product layer, and the seven-state model reproduces SEER incidence with the module at L3. The tissue-layer chain from L1 to L4 is supported; L4 to L5 is the cell-mutation axiom, and L5 to L7 is the established mutation-clone chain. The five-action module is the tissue-layer target for lung cancer.

References

1. National Cancer Institute. SEER Cancer Stat Facts: Lung and Bronchus Cancer.

2. Yoshida K, Gowers KHC, Lee-Six H, et al. Tobacco smoking and somatic mutations in human bronchial epithelium. Nature. 2020;578(7794):266-272. PMID: 31996850.

3. Ishizumi T, McWilliams A, MacAulay C, et al. Natural history of bronchial preinvasive lesions. Cancer Metastasis Rev. 2010;29(1):5-14. PMID: 20112052.

4. Fu F, Shang J, Yan Y, et al. Genomic and transcriptomic dynamics in the stepwise progression of lung adenocarcinoma. Cell Res. 2025;35(12):1037-1055. PMID: 41345544.

5. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender. Cancer Epidemiol Biomarkers Prev. 2015;24(1):154-161. PMID: 25312998.

6. McMahon PM, Kong CY, Johnson BE, et al. The MGH-HMS lung cancer policy model. Risk Anal. 2012;32 Suppl 1:S117-S124. PMID: 22882882.

7. Chrysanthopoulou SA, Rutter CM, Gatsonis CA. Bayesian versus empirical calibration of microsimulation models. Med Decis Making. 2021;41(6):714-726. PMID: 33966518.

8. Nargund RS, Ishizawa S, Eghbalizarch M, et al. Natural history models for lung cancer: a scoping review. Lung Cancer. 2025;203:108495. PMID: 40174386.

9. Ferrucci L, Donega S, Picca A, et al. Sarcopenia as a manifestation of systemic energetic failure with ageing. Lancet Healthy Longev. 2026;7(9):100904. PMID: 42805229.

10. Richardson RB, Anghel CV, Deng DS. Profound synchrony of age-specific incidence rates and tumor suppression. Aging (Albany NY). 2021;13(20):23545-23578. PMID: 34695806.

11. Rosales BM, De La Mata N, Vajdic CM, et al. Cancer mortality in people receiving dialysis for kidney failure. Am J Kidney Dis. 2022;80(4):449-461. PMID: 35500725.

12. Arakawa Y, Ishida Y, Higa A, et al. Cancer risk in haemodialysis patients. Nephrology (Carlton). 2026;31(10):e70293. PMID: 42865121.

13. National Cancer Institute. SEER Cancer Statistics Review 1975-2016, Table 15.10.

14. Peto R, Darby S, Deo H, et al. Smoking, smoking cessation, and lung cancer in the UK since 1950. BMJ. 2000;321(7257):323-329. PMID: 10926586.

15. Moro-Sibilot D, Fievet F, Jeanmart M, et al. Clinical prognostic indicators of high-grade pre-invasive bronchial lesions. Eur Respir J. 2004;24(1):24-29. PMID: 15293601.

16. Kakinuma R, Noguchi M, Ashizawa K, et al. Natural history of pulmonary subsolid nodules. J Thorac Oncol. 2016;11(7):1012-1028. PMID: 27089851.

17. Rodriguez de Vera-Gomez P, Galindo-Gallardo M, Gonzalez-Pacheco M, et al. Malnutrition severity, morphofunctional assessment, and survival in patients with lung cancer. Lung. 2026;204(1). PMID: 42595921.

18. Lu Z, Fu S, Li W, et al. Prognostic role of C-reactive protein to albumin ratio in lung cancer. Chronic Dis Transl Med. 2024;10(1):31-39. PMID: 38450307.

19. Zhao S, Jiang T, Zhang L, et al. Clinicopathological and prognostic significance of regulatory T cells in non-small cell lung cancer. Oncotarget. 2016;7(24):36065-36073. PMID: 27153545.

20. Sheida F, Alipour N, Naseri A, et al. Prognostic value of myeloid-derived suppressor cells in lung cancer. Expert Rev Anticancer Ther. 2026;26(3):345-360. PMID: 41195653.

21. Liao C, Yu Z, Guo W, et al. Prognostic value of circulating inflammatory factors in non-small cell lung cancer. Cancer Biomark. 2014;14(6):469-481. PMID: 25335739.

22. Pan T, Duan R, Xu Z, et al. GDF-15 as a biomarker for diagnosis and prognosis of lung cancer. Front Oncol. 2025;15:1447990. PMID: 40144214.

23. Fleming JS, Conway J, Bennett MJ, et al. Quantitative assessment of mucociliary clearance in smokers with mild-to-moderate COPD and chronic bronchitis. J Aerosol Med Pulm Drug Deliv. 2019;32(4):175-188. PMID: 30848685.

24. Pincikova T, Merikallio H, Kotortsi I, et al. Expression levels of MUC5AC and MUC5B in airway goblet cells. Int J Mol Sci. 2024;25(24):13653. PMID: 39769414.

25. Zhang Z, Li Y, Yan X, et al. Pretreatment lactate dehydrogenase may predict outcome of advanced non-small-cell lung cancer. Cancer Med. 2019;8(4):1467-1473. PMID: 30848091.

26. Cao L, Zhong J, Liu Z, et al. Increased LOXL2 expression is related to poor prognosis in lung squamous cell carcinoma. J Thorac Dis. 2024;16(1):581-592. PMID: 38410543.

27. Picchi MA, Thomas CL, Grimes MJ, et al. Reduced DNA repair capacity by lymphocytes predicts risk for lung cancer and COPD. Environ Res. 2026;303(Pt 2):124787. PMID: 42150679.

28. Janik J, Swoboda M, Janowska B, et al. 8-Oxoguanine incision activity is impaired in lung tissues of NSCLC patients. Mutat Res. 2011;709-710:21-31. PMID: 21376741.

29. Smesseim I, van Boerdonk RA, Dickhoff C, et al. Focal FDG uptake predicts progression of pre-invasive squamous bronchial lesions. Thorac Cancer. 2023;14(9):840-847. PMID: 36802171.

30. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395-409. PMID: 21714641.

31. Short KR, Bigelow ML, Kahl J, et al. Decline in skeletal muscle mitochondrial function with aging in humans. Proc Natl Acad Sci U S A. 2005;102(15):5618-5623. PMID: 15800038.

32. Gay EL, Coen PM, Harrison S, et al. Sex differences in the association between skeletal muscle energetics and perceived physical fatigability. Geroscience. 2025;47(2):1999-2013. PMID: 39436549.

33. Brodeau C, Joly C, Chekroun A, et al. Identification of functional cellular markers related to human health, frailty and chronological age. Aging Cell. 2025;24(9):e70153. PMID: 40591484.

34. Kawai E, Nakanishi J, Kumazawa N, et al. Skin surface electric potential as an indicator of skin condition. Exp Dermatol. 2008;17(8):688-692. PMID: 18341574.

35. Denda M, Ashida Y, Inoue K, et al. Skin surface electric potential induced by ion-flux through epidermal cell layers. Biochem Biophys Res Commun. 2001;284(1):112-117. PMID: 11374879.

36. Nuccitelli R, Nuccitelli P, Li C, et al. The electric field near human skin wounds declines with age. Wound Repair Regen. 2011;19(5):645-655. PMID: 22092802.

37. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med. 2020;382(6):503-513. PMID: 31995683.

38. van Boerdonk RA, Smesseim I, Heideman DA, et al. Close surveillance with long-term follow-up of subjects with preinvasive endobronchial lesions. Am J Respir Crit Care Med. 2015;192(12):1483-1489. PMID: 26275031.

39. Ma Z, Wang Z, Li Y, et al. Detection and treatment of lung adenocarcinoma at pre-/minimally invasive stage. J Cancer Res Clin Oncol. 2022;148(10):2717-2722. PMID: 35524781.

40. Jonsson S, Franklin WA, Varella-Garcia M, et al. Prevalence, molecular markers, and outcome of bronchial squamous carcinoma in situ. APMIS. 2023;131(10):513-527. PMID: 37608782.

41. Nie M, Yao K, Zhu X, et al. Evolutionary metabolic landscape from preneoplasia to invasive lung adenocarcinoma. Nat Commun. 2021;12(1):6479. PMID: 34759281.

42. Xu X, Gong T, Huo Z, et al. Proteomic landscape of lung adenocarcinoma precancerous lesions reveals mitochondrial dsRNA-associated immune and stress responses. Am J Cancer Res. 2025;15(9):3888-3903. PMID: 41113990.

43. Mascaux C, Angelova M, Vasaturo A, et al. Immune evasion before tumour invasion in early lung squamous carcinogenesis. Nature. 2019;571(7766):570-575. PMID: 31243362.

44. Ning B, Tilston-Lunel AM, Simonetti J, et al. Convergence of YAP/TAZ, TEAD and TP63 activity is associated with bronchial premalignant severity and progression. J Exp Clin Cancer Res. 2023;42(1):116. PMID: 37150829.

45. Fleming JS, Conway J, Bennett MJ, et al. Quantitative assessment of regional mucociliary clearance in smokers with mild-to-moderate COPD and chronic bronchitis. J Aerosol Med Pulm Drug Deliv. 2020;33(6):342-356. PMID: 32640859.

46. Bruse S, Petersen H, Weissfeld J, et al. Increased methylation of lung cancer-associated genes in sputum DNA of former smokers with chronic mucous hypersecretion. Respir Res. 2014;15(1):2. PMID: 24405663.

47. Oo MW, Hikita T, Mashima T, et al. Cancer-associated fibroblast-derived SOD3 enhances lymphangiogenesis to drive metastasis in lung adenocarcinoma. Angiogenesis. 2025;28(4):51. PMID: 41028519.

48. Chen C, Feng Y, Ko FCF, et al. Lung cancer-associated fibroblasts-mediated collagen deposition drives mediastinal lymph node metastasis in non-small cell lung cancer. Front Oncol. 2025;15:1597585. PMID: 40575163.

49. Lantuejoul S, Raynaud C, Salameire D, et al. Telomere maintenance and DNA damage responses during lung carcinogenesis. Clin Cancer Res. 2010;16(11):2979-2988. PMID: 20404006.

50. Lan L, Wang X, Jiang X, et al. Spatiotemporal analysis of molecular continuity in pathogenesis of squamous cell carcinoma. Cancer Commun (Lond). 2026;46:0051. PMID: 42898679.

51. Hong JH, Park S, Kim H, et al. Volume and mass doubling time of lung adenocarcinoma according to WHO histologic classification. Korean J Radiol. 2021;22(3):464-475. PMID: 33169551.

52. Tang EK, Wu YJ, Chen CS, et al. Prediction of the stage shift growth of early-stage lung adenocarcinomas by volume-doubling time. Quant Imaging Med Surg. 2024;14(6):3983-3996. PMID: 38846271.

53. SHARMA R, KHUBCHANDANI J. Global, regional, and national burden of tracheal, bronchus, and lung cancer in 2022: evidence from the GLOBOCAN study[J]. Epidemiologia (Basel), 2024, 5(4): 785-795. DOI: 10.3390/epidemiologia5040053.

54. SOTO A M, SONNENSCHEIN C. The somatic mutation theory of cancer: growing problems with the paradigm?[J]. Bioessays, 2004, 26(10): 1097-1107. DOI: 10.1002/bies.20087.

55. VENDRAMIN R, LITCHFIELD K, SWANTON C. Cancer evolution: Darwin and beyond[J]. EMBO J, 2021, 40(18): e108389. DOI: 10.15252/embj.2021108389.

56. SCHREIBER R D, OLD L J, SMYTH M J. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion[J]. Science, 2011, 331(6024): 1565-1570. DOI: 10.1126/science.1203486.

57. RICHARDSON R B, ANGHEL C V, DENG D S. Profound synchrony of age-specific incidence rates and tumor suppression for different cancer types as revealed by the multistage-senescence model of carcinogenesis[J]. Aging (Albany NY), 2021, 13(20): 23545-23578. DOI: 10.18632/aging.203651.

58. MATAS J, KOHRN B, FREDRICKSON J, et al. Colorectal cancer is associated with the presence of cancer driver mutations in normal colon[J]. Cancer Res, 2022, 82(8): 1492-1502. DOI: 10.1158/0008-5472.CAN-21-3607.

59. MURAI K, DENTRO S, ONG S H, et al. p53 mutation in normal esophagus promotes multiple stages of carcinogenesis but is constrained by clonal competition[J]. Nat Commun, 2022, 13(1): 6206. DOI: 10.1038/s41467-022-33945-y.

60. PAYNE S L, RAM P, SRINIVASAN D H, et al. Potassium channel-driven bioelectric signalling regulates metastasis in triple-negative breast cancer[J]. EBioMedicine, 2022, 75: 103767. DOI: 10.1016/j.ebiom.2021.103767.

61. KOFMAN K, LEVIN M. Bioelectric pharmacology of cancer: a systematic review of ion channel drugs affecting the cancer phenotype[J]. Prog Biophys Mol Biol, 2024, 191: 25-39. DOI: 10.1016/j.pbiomolbio.2024.07.005.

62. GAJOVIC N, JURISEVIC M, PANTIC J, et al. Attenuation of NK cells facilitates mammary tumor growth in streptozotocin-induced diabetes in mice[J]. Endocr Relat Cancer, 2018, 25(4): 493-507. DOI: 10.1530/ERC-17-0529.

63. SURIBEN R, CHEN M, HIGBEE J, et al. Antibody-mediated inhibition of GDF15-GFRAL activity reverses cancer cachexia in mice[J]. Nat Med, 2020, 26(8): 1264-1270. DOI: 10.1038/s41591-020-0945-x.

64. ZHU X A, STAROSTA S, FERRER M, et al. A neuroimmune circuit mediates cancer cachexia-associated apathy[J]. Science, 2025, 388(6743): eadm8857. DOI: 10.1126/science.adm8857.

65. TRAC P T, THAI T L, LINCK V, et al. Alveolar nonselective channels are ASIC1a/α-ENaC channels and contribute to AFC[J]. Am J Physiol Lung Cell Mol Physiol, 2017, 312(6): L797-L811. DOI: 10.1152/ajplung.00379.2016.

66. GILLE T, RANDRIANARISON-PELLAN N, GOOLAERTS A, et al. Hypoxia-induced inhibition of epithelial Na+ channels in the lung: role of Nedd4-2 and the ubiquitin-proteasome pathway[J]. Am J Respir Cell Mol Biol, 2014, 50(3): 526-537. DOI: 10.1165/rcmb.2012-0518OC.

67. VADÁSZ I, MORTY R E, OLSCHEWSKI A, et al. Thrombin impairs alveolar fluid clearance by promoting endocytosis of Na+,K+-ATPase[J]. Am J Respir Cell Mol Biol, 2005, 33(4): 343-354. DOI: 10.1165/rcmb.2004-0407OC.

68. SONG Y, FUKUDA N, BAI C, et al. Role of aquaporins in alveolar fluid clearance in neonatal and adult lung, and in oedema formation following acute lung injury: studies in transgenic aquaporin null mice[J]. J Physiol, 2000, 525(Pt 3): 771-779. DOI: 10.1111/j.1469-7793.2000.00771.x.

69. SRINIVASAN R S, ESCOBEDO N, YANG Y, et al. The Prox1-Vegfr3 feedback loop maintains the identity and the number of lymphatic endothelial cell progenitors[J]. Genes Dev, 2014, 28(19): 2175-2187. DOI: 10.1101/gad.216226.113.

70. MA W, GIL H J, LIU X, et al. Mitochondrial respiration controls the Prox1-Vegfr3 feedback loop during lymphatic endothelial cell fate specification and maintenance[J]. Sci Adv, 2021, 7(18): eabe7359. DOI: 10.1126/sciadv.abe7359.

71. SABINE A, AGALAROV Y, MABY-EL HAJJAMI H, et al. Mechanotransduction, PROX1, and FOXC2 cooperate to control connexin37 and calcineurin during lymphatic-valve formation[J]. Dev Cell, 2012, 22(2): 430-445. DOI: 10.1016/j.devcel.2011.12.020.

72. CHAKRABORTY S, SINGH A, WANG L, et al. Trained immunity of alveolar macrophages enhances injury resolution via KLF4-MERTK-mediated efferocytosis[J]. J Exp Med, 2023, 220(11): e20221388. DOI: 10.1084/jem.20221388.

73. JANG M, AN J, OH S W, et al. Matrix stiffness epigenetically regulates the oncogenic activation of the Yes-associated protein in gastric cancer[J]. Nat Biomed Eng, 2021, 5(1): 114-123. DOI: 10.1038/s41551-020-00657-x.

74. MOKHTARI R B, SAMPATH D, EVERSOLE P, et al. An Agrin-YAP/TAZ rigidity sensing module drives EGFR-addicted lung tumorigenesis[J]. Adv Sci (Weinh), 2025, 12(20): e2413443. DOI: 10.1002/advs.202413443.

75. CHENG T, LIU Q, ZHANG R, et al. Lysyl oxidase promotes bleomycin-induced lung fibrosis through modulating inflammation[J]. J Mol Cell Biol, 2014, 6(6): 506-515. DOI: 10.1093/jmcb/mju039.

76. COLEGIO O R, CHU N Q, SZABO A L, et al. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid[J]. Nature, 2014, 513(7519): 559-563. DOI: 10.1038/nature13490.

77. PRIMEAU A J, RENDON A, HEDLEY D, et al. The distribution of the anticancer drug doxorubicin in relation to blood vessels in solid tumors[J]. Clin Cancer Res, 2005, 11(24 Pt 1): 8782-8788. DOI: 10.1158/1078-0432.CCR-05-1664.

78. LINDAHL T. Instability and decay of the primary structure of DNA[J]. Nature, 1993, 362(6422): 709-715. DOI: 10.1038/362709a0.

79. ARAI T, KELLY V P, KOMORO K, et al. Cell proliferation in liver of Mmh/Ogg1-deficient mice enhances mutation frequency because of the presence of 8-hydroxyguanine in DNA[J]. Cancer Res, 2003, 63(14): 4287-4292.

80. FOUQUEREL E, SOBOL R W. ARTD1 (PARP1) activation and NAD+ in DNA repair and cell death[J]. DNA Repair (Amst), 2014, 23: 27-32. DOI: 10.1016/j.dnarep.2014.08.004.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTablesS1S7.docx