



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

의학박사 학위논문

수술적 절제를 시행 받은 비소세포성

폐암 환자에서

클론성 조혈증의 임상적 유의성

Clinical significance of clonal hematopoiesis

in patients with surgically resected non-small cell lung  
cancer

울산대학교 대학원

의학과

윤재광

수술적 절제를 시행 받은 비소세포성  
폐암 환자에서  
클론성 조혈증의 임상적 유의성

지도교수     최세훈

이 논문을 의학박사학위 논문으로 제출함

2021 년 8 월

울산대학교 대학원

의 학 과

윤재광

윤재광의 의학박사학위 논문을 인준함

|      |       |   |
|------|-------|---|
| 심사위원 | 김 동 관 | 인 |
| 심사위원 | 최 세 훈 | 인 |
| 심사위원 | 김 규 표 | 인 |
| 심사위원 | 황 희 상 | 인 |
| 심사위원 | 고 영 일 | 인 |

울 산 대 학 교 대 학 원

2021 년 8 월

## **Abstract**

**Background:** Clonal hematopoiesis (CH), somatic mutations derived from the expansion of clonal populations of blood cells, is associated with aging, increased risk of hematologic malignancies and cardiovascular disease, and shorter overall survival. Recent studies suggest CH is more prevalent in patients with solid cancer, with approximately 30% harboring CH mutations in their blood. We evaluated the clinical impact of preoperative CH on the survival outcomes of patients with non-small cell lung cancer (NSCLC) who underwent surgical resection followed by adjuvant therapy.

**Methods:** We retrospectively reviewed the medical records and analyzed the blood samples of 341 consecutive patients with NSCLC who underwent surgery followed by adjuvant therapy from 2013 to 2017. We analyzed 89 genes found to be associated with CH by targeted deep sequencing of blood samples from 341 individuals collected before surgery to assess the clinical relevance of CH mutations. To minimize the possible selection bias between the two groups, a propensity score matching (PSM) technique was adopted.

**Results:** We identified CH in 23% (77/341) of patients with NSCLC. Patients with CH mutations had worse overall survival (OS) than those without (5-year OS rate: 42.8% vs. 59.8%,  $p = 0.005$ ), which remained the same after PSM (5-year rate: 43.2% vs. 58.7%,  $p = 0.019$ ). According to the cause of death, the presence of CH was associated with an increase in lung cancer mortality ( $p = 0.017$ ), especially without evident cancer progression ( $p < 0.001$ ). In multivariable analysis, the presence of CH, along with histologic type and overall stage was a significant prognostic factor for OS in patients with advanced NSCLC who underwent adjuvant therapy (hazard ratio [95% confidence interval] = 1.60 [1.12–2.30],  $p = 0.011$ ). Age, which was significant in univariable analysis, became insignificant after the adjustment with several covariates, including the presence of CH.

**Conclusions:** In resected NSCLC, preoperative CH mutations might amplify CH-related adverse outcomes through adjuvant treatments, resulting in poor survival results.

Abstract word count: 317

## Table of Contents

|  |           |
|--|-----------|
| <i>Abstract</i> .....  | <i>iv</i> |
| <i>Table of Contents</i> .....   | <i>v</i>  |
| <i>Introduction</i> .....  | <i>1</i>  |
| <i>Methods</i> .....   | <i>1</i>  |
| <b>Patients</b> .....  | <b>1</b>  |
| <b>Sample Processing and Sequencing</b> .....  | <b>2</b>  |
| <b>Preoperative and Postoperative Management</b> .....   | <b>2</b>  |
| <b>Definitions</b> .....   | <b>3</b>  |
| <b>Statistical analysis</b> .....  | <b>3</b>  |
| <i>Results</i> .....   | <i>4</i>  |
| <b>Characteristics of CH</b> .....   | <b>4</b>  |
| <b>Patient Characteristics</b> .....   | <b>5</b>  |
| <b>Survival Analysis</b> .....   | <b>5</b>  |
| <i>Discussion</i> .....  | <i>6</i>  |
| <i>Conclusion</i> .....  | <i>9</i>  |
| <i>References</i> .....  | <i>10</i> |
| <i>Table 1. Baseline characteristics of enrolled patients according to the presence of clonal hematopoiesis</i> .....  | <i>12</i> |
| <i>Table 2. Sensitivity test for clonal hematopoiesis, clonal hematopoiesis-potential driver, and clonal hematopoiesis non-potential driver</i> .....  | <i>14</i> |
| <i>Table 3. Univariable and multivariable analysis for overall survival of all patients</i> .....  | <i>15</i> |
| <i>Figure 1. CONSORT diagram</i> .....   | <i>16</i> |
| <i>Figure 2. Overall survival (A) and recurrence-free survival (B) of patients following the presence of CH mutations in the entire cohort. Overall survival (C) of patients following the presence of CH mutations after propensity score matching</i> .....  | <i>17</i> |
| <i>Figure 3. (A) Cumulative non-lung cancer mortality according to the presence of CH. (B) Cumulative lung cancer mortality according to the presence of CH. (C) Cumulative lung cancer mortality with evident cancer progression according to the presence of CH. (D) Cumulative lung cancer mortality without evident cancer progression according to the presence of CH</i> ..... | <i>18</i> |

|  |    |
|--|----|
| <i>Supplementary Methods</i> .....   | 19 |
| <i>Supplementary Table 1. Eighty-nine clonal hematopoiesis genes targeted in the NGS panel.</i> .....  | 21 |
| <i>Supplementary Table 2. Sensitivity test for clonal hematopoiesis based on the variant allele fraction value.</i> .....  | 22 |
| <i>Supplementary Table 3. Baseline characteristics of enrolled patients according to the presence of clonal hematopoiesis after propensity score matching.</i> .....   | 23 |
| <i>Supplementary Table 4. Cause of deaths for the entire observation period.</i> .....   | 24 |
| <i>Supplementary Figure 1. Characteristics of CH mutations identified in the overall cohort. (A) CH association with age, and age distribution of patients in the cohort. (B) Number of mutations harbored per patient. (C) The contribution of individual genes to the total number of potential driver somatic mutations that were observed.</i> ..... | 25 |
| <i>Supplementary Figure 2. Overall survival of patients following the presence of CH mutations in (A) stage IIB, (B) stage IIIA, and (C) stage IIIB.</i> .....   | 26 |
| <i>Korean abstracts (국문요약)</i> .....   | 27 |

## **Introduction**

Clonal hematopoiesis (CH) is a condition defined by the expansion of clonally derived hematopoietic stem cells (HSCs) that harbor somatic mutations in leukemia-associated genes, which can be detected by next-generation sequencing (NGS).<sup>1-3</sup> CH is associated with aging, and can be found without hematological malignancies (hence originally referred to as clonal hematopoiesis of indeterminate potential [CHIP]).<sup>4,5</sup> In addition, it has a significant association with tobacco use, prior radiation therapy (RTx), and/or prior exposure to chemotherapy (CTx).<sup>6</sup>

CH is reported to increase the incidence of subsequent cardiovascular diseases and hematological malignancies.<sup>1,4,7</sup> CH causes cardiovascular disease as a result of mutated genes amplifying the innate inflammatory response, a known contributing factor for developing atherosclerosis.<sup>3</sup> Indeed, the clinical significance of CH in various disease has recently started to gain attraction.<sup>8-10</sup>

Of particular interest is the impact of CH on cancer survivors who have previously undergone cancer-related therapy. A recent study demonstrated that CH is common in patients with solid tumors.<sup>6</sup> Considering the altered immune response by CH, CH may also alter the clinical consequences of the cancer. Hence, an in-depth study into how CH might influence cancer recurrence and response to therapy will help to decide the surveillance protocol, such as screening, follow-up duration, and risk-directed therapeutic approaches for high-priority groups.

Lung cancer is the most commonly diagnosed cancer and the leading cause of cancer-related deaths worldwide.<sup>11</sup> Given the critical role of inflammation in the pathogenesis of lung cancer,<sup>12</sup> there is a possibility that CH influences the prognosis in patients with non-small cell lung cancer (NSCLC). Although several prognostic factors, such as age, sex, and cancer stage, have been identified,<sup>13</sup> further study to determine novel factors in the era of NGS is highly encouraged.

In this study, we evaluated the clinical impact of preoperatively existing CH on recurrence and survival in patients with NSCLC who received surgical resection followed by adjuvant therapy.

## **Methods**

### **Patients**

All clinical records of patients who underwent surgery for NSCLC between January 2013 and December 2017 were reviewed from the lung cancer database of Asan Medical Center, Seoul, Korea. We included patients who received adjuvant CTx or chemoradiation therapy (CRTx) for pathological stage II or III NSCLC (Figure 1). The exclusion criteria were as follows: i) patients with previous or current malignancy other than lung cancer; ii) patients who received neoadjuvant therapy; iii) patients

who underwent sublobar resection (wedge resection or segmentectomy); iv) patients with incomplete resection, and v) patients who died within 30 days after surgery. After exclusion according to the criteria, we identified 350 patients with blood samples that were collected before surgery and stored in Asan Bio-Resource Center, Korea Biobank Network. Of these, 9 samples were excluded due to sample degradation, and 341 patients with blood samples were finally enrolled. The study was approved by the Institutional Review Board of Asan Medical Center (2020–0906). All participants provided written informed consent.

### **Sample Processing and Sequencing**

Blood-derived DNA from a patient was used for targeted NGS with a custom panel containing 89 genes. The sequencing libraries were prepared following the SureSelect XT HS Target Enrichment System (Agilent, Santa Clara, CA) protocol. The libraries were sequenced on the Illumina NovaSeq6000 platform (Illumina, San Diego, CA) with 150 bp paired-end following the manufacturer's protocols. The mean depth of coverage of an analysis ready BAM was more than 800× (See supplementary methods for details).

### **Preoperative and Postoperative Management**

Patient workup for diagnosis, staging, and surgical resection was performed according to well-established, widely accepted protocols, the details of which are described elsewhere.<sup>14</sup> When clinical N2 (cN2) disease was suspected on computed tomography (CT) or positron emission tomography (PET), mediastinal LN biopsy was performed for suspicious nodes using endobronchial ultrasonography (EBUS), mediastinoscopy, or endoscopic ultrasound (EUS). Treatment plans for biopsy-proven N2 disease were determined by a multidisciplinary team, including medical oncologists, radiologists, and thoracic surgeons. Patients in the study sample were retrospectively staged according to the American Joint Committee on Cancer (AJCC) 8<sup>th</sup> edition.<sup>15</sup>

Adjuvant CTx was recommended for all patients with stage II and III, except when the patient was > 75 years of age or in poor physical condition, according to the judgment of the multidisciplinary team. Systemic CTx with a platinum-based regimen was recommended for 4–6 weeks after surgery, with a total of four cycles of treatment. From 2008, when the use of targeted therapy became common for patients with activating mutations in the epidermal growth factor receptor (EGFR), a tyrosine kinase inhibitor was mainly used when recurrence occurred after the first adjuvant CTx. For adjuvant RTx for stage III, a daily dose of 1.8 Gy was administered up to a total dose of 50.4 Gy for patients who underwent complete resection or 55–60 Gy for patients with positive resection margins. Among patients

who underwent complete resection, a considerable number of patients with a single N2 node metastasis skipped adjuvant RTx.

Follow-up information for all patients was obtained through clinic follow-up notes every 6 months during the first 2 years after surgery, and every year thereafter. Chest CT scans were performed at every clinical visit or at any time when disease recurrence was suspected. Treatment modalities and chemotherapeutic regimens for relapse cases were determined at the discretion of the attending physician.

## **Definitions**

Eighty-nine genes frequently detected in CH were custom selected and a variant allele frequency (VAF) of  $\geq 2\%$  was set as the cut-off for carrier (See Supplementary Table 1 for gene list).<sup>1,2,4,7,16,17</sup> The following criteria were used to classify potential driver (PD) mutations: 1) any somatic variant in COSMIC occurring in the “hematopoietic and lymphoid” category more than once, or in the categories more than ten times; and 2) any somatic variant with loss of function effects (frame shift, stop gain, mutations at splice site donor/acceptor).<sup>6</sup> CH cases without PD mutations were classified as non-PD cases.

Overall survival (OS) was defined as the time interval between the date of operation and the date of death, which was determined by reviewing the records from the Korean National Security Death Index Database. Non-lung cancer mortality was defined as death with a known cause not due to lung cancer and without previous cancer recurrence. Lung cancer mortality included all other deaths, that is, 1) deaths resulting from evident tumor progression, 2) deaths from other causes after cancer recurrence, and 3) deaths from an unknown cause. For the accessibility of the study, we subdivided lung cancer mortality into two categories: 1) those with evident cancer progression, and 2) those without evident cancer progression (deaths from other cause after cancer recurrence or unknown origin). Recurrence-free survival (RFS) was calculated as the time between the date of resection and the date of recurrence, and patients without recurrence were censored at the latest timepoint known to be recurrence-free.

## **Statistical analysis**

Continuous variables are presented as means and standard deviations, and categorical variables as count and percentage. The normality of individual parameter distributions was assessed with the Shapiro–Wilk test. Student’s t-test or Wilcoxon rank–sum test was used to compare the two groups in terms of continuous variables, and the chi-square test or Fisher’s exact test was applied for categorical variables. After propensity score matching (PSM), the McNemar’s test and paired t-tests were used to analyze the propensity score-matched pairs. The OS and RFS outcomes were defined using Kaplan–

Meier curves. The differences in the survival rates were analyzed using the log-rank test. The Cox proportional hazards model was used for univariable and multivariable analyses to identify the clinical impact of CH on survival outcomes. We developed two types of multivariable Cox model: age and sex were involved as covariates to adjust for the sensitivity test on CH (Table 2), and forward stepwise selection was used as the selection procedure for the likelihood ratio test ( $p\text{-value} \leq 0.10$  for entering the model and  $p\text{-value} \leq 0.05$  for staying in the model) to identify the prognostic effect of CH in patients who received adjuvant therapy (Table 3). The proportional hazards assumption for the Cox regression models was tested with Schoenfeld residuals.

PSM was applied to adjust for the possible selection bias derived from a retrospective nonrandomized cohort for generating the two groups (CH-positive and CH-negative) with similar characteristics. Eleven variables (Table 1) were used to balance the clinical characteristics of the two groups.<sup>18</sup> For PSM, observation pairs with equivalent propensity scores were selected with nearest-neighbor matching and a caliper width of 0.25 of the standard deviation. CH-negative patients were randomly matched to CH-positive patients at a ratio of 1:1. Balance between the groups was assessed using standardized mean differences (SMDs). An absolute standardized difference of  $\leq 0.1$  was considered to indicate the ideal balance and that of  $\leq 0.2$  was considered to indicate acceptable balance.<sup>19</sup>

All statistical calculations were performed using R version 4.0.2 (The R Foundation for Statistical Computing, Vienna, Austria) using the “Survival,” “MatchIt,” “dplyr,” “sad,” “ggplot2,” “GGally,” “survminer,” and “rms” packages. All reported p-values are two-sided. P-values  $< 0.05$  were considered significant.

## **Results**

### **Characteristics of CH**

The mean age of the patients at the time of sample collection was  $60.2 \pm 8.6$  years. Of the total 341 patients, CH was found in 77 (22.6%) patients. The prevalence of CH was 4.8%, 16.9%, 25.2%, and 42.6% in patients in their 40s, 50s, 60s, and 70s, respectively, showing a continuous increase with age (Supplementary Figure 1A). As for the number of mutations, single mutation was the most common in 81.8% of patients, two mutations in 14.3%, and three mutations in 3.9% (Supplementary Figure 1B). Mutations in DNMT3A (33.0%) were the most common, followed by ASXL1 (11.7%), TET2 (8.5%), and PPM1D (8.5%); these four genes accounted for 61.7% of all mutations detected (Supplementary Figure 1C). A sensitivity test based on several cut-off values for VAF showed that when the cutoff

value of VAF was  $\geq 1.8\%$ , the p-values were the lowest in the univariable and multivariable (adjusted by age and sex) Cox analysis for OS, at 0.003 and 0.035, respectively (Supplementary Table 2). However, a cutoff value of VAF  $\geq 2\%$  had similar outcomes to that  $\geq 1.8\%$  (p-value = 0.005 and 0.055 for univariable and multivariable analysis, respectively) (Supplementary Table 2).

The prevalence of CH-PD and CH non-PD in the study patients was 17.0% (58/341) and 7.9% (27/341), respectively, and 75.3% (58/77) and 35.1% (27/77) in patients with CH mutations. A sensitivity test was also performed in these patients. As a result, CH-PD was still observed as a significant factor after the adjustment by age and sex in multivariable analysis (p = 0.038), whereas CH non-PD was not significant in univariable analysis (p = 0.157) (Table 2).

### **Patient Characteristics**

The mean postoperative follow-up duration was  $51.5 \pm 23.9$  months. The baseline demographics of the patients and tumor characteristics are listed in Table 1. Generally, patients with CH (n = 77) were older (p < 0.001) and had a higher rate of males (p = 0.009) than those without (n = 264). There were no significant differences in smoking history (p = 0.144), the number of comorbidities (p = 0.554), the rate of EGFR mutation (p = 0.452), the distribution of histology (p = 0.720) and overall stage (p = 0.367), and the type of adjuvant therapy (p = 0.298) between the two groups. After PSM, all variables, including age (p = 0.604) and sex (p = 0.833), became similar between the two groups and were well balanced (all SMDs < 0.2) (Supplementary Table 3).

### **Survival Analysis**

Overall, 42 patients with CH (n = 77) and 105 patients without CH (n = 264) had died by the end of follow-up, and their 5-year OS rates were 42.8% and 59.8%, respectively. Recurrence events occurred in 43 and 137 patients with and without CH, respectively, and their 5-year RFS rates were 40.0% and 44.0%, respectively. Detailed information for the cause of death is summarized in Supplementary Table 3.

The Kaplan–Meier survival curves according to the presence of CH are plotted in Figure 2. While there was no significant difference in RFS between the two groups (p = 0.662), patients with CH had worse OS than those without CH (p = 0.005) (Figures 2A and 2B). After PSM, patients with CH still had a significantly worse survival rate than those without CH (p = 0.019) (Figure 2C).

According to the cause of death, non-lung cancer mortality was similar regardless of CH (p = 0.317), but lung cancer mortality was significantly higher in patients with CH than those without (p = 0.017) (Figures 3A and 3B). Among patients with lung cancer mortality, there was no significant

difference in mortality due to cancer progression ( $p = 0.651$ ), but a significant difference was shown in mortality without evident cancer progression ( $p < 0.001$ ) (Figures 3C and 3D).

In multivariable Cox analysis, the presence of CH, along with histologic type and overall stage was a significant prognostic factor for OS in patients with advanced NSCLC who underwent adjuvant therapy (hazard ratio [HR] [95% confidence interval] = 1.60 [1.12–2.30],  $p = 0.011$ ) (Table 3). Age and the number of comorbidities, which were significant in univariable analysis, became insignificant after the adjustment with several covariates, including the presence of CH. EGFR mutation was not a significant factor in univariable analysis (HR [95% confidence interval] = 1.10 [0.76–1.58],  $p = 0.621$ ).

## Discussion

In this study, we examined the prevalence and the traits for CH in patients with advanced NSCLC. Further, the clinical impact of preoperatively existing CH on survival outcomes was evaluated in overall patients and after PSM. As a result, 23% of the patients had CH before surgery, which was increased with aging. Mutation in DNMT3A was the most common, followed by ASXL1, TET2, and PPM1D. Additionally, CH-PD accounted for three quarters of CH and acted as a more significant prognostic factor for OS than CH itself. The presence of CH before surgery was significantly associated with an increase in overall mortality, lung cancer mortality, and without evident cancer progression. Multivariable Cox analysis revealed that the presence of CH was a significant prognostic factor among patients who underwent surgery followed by adjuvant therapy for advanced stage NSCLC. The prognostic effect of CH was the same after adopting a rigorous risk-adjustment methodology to properly adjust the baseline covariates between the two groups.

Since the first detection of non-random X-chromosome inactivation in healthy elderly women,<sup>20</sup> research into CH and our understanding has grown considerably over the past few years.<sup>1,2,4,5,7</sup> Several studies have demonstrated that CH, a common age-dependent state, is closely related to an increased risk of subsequent hematologic malignancies, cardiovascular events, and adverse outcomes in patients with advanced solid tumors.<sup>3,10</sup> Patients with cancer have higher rates of CH than healthy individuals, and CH is associated with shorter patient survival.<sup>10</sup> This is probably due to a high genetic predisposition to malignancy, prolonged exposure to a carcinogenic environment, and cancer-related therapy with genotoxic therapies.<sup>2</sup> Especially in cancer-related therapy, CH has been reported to have a significant association with CTx and RTx, which means the local and systemic treatment might promote clonal outgrowth of HSCs.<sup>10</sup> From this perspective, we hypothesized that preoperatively existing CH amplifies the series of processes that trigger CH-related adverse outcomes through cancer-related treatments, resulting in poor survival results.

According to our survival analysis, the presence of CH was significantly associated with poor

OS ( $p = 0.005$ ) (Figure 2A); however, given the positive correlation between CH and age, it is inappropriate to determine the effect of CH merely with this result. To overcome this problem, we conducted two types of statistical adjustment. After the adjustment of age and sex with multivariable analysis, the prognostic effect of CH was marginally significant ( $p = 0.055$ ), which is expected to be evident when analyzed in a large number of patients (Supplementary Table 2). Additionally, the fact that CH, and not age, was included as the final prognostic factor through forward stepwise selection in multivariable analysis suggests that CH is more closely related to survival than age (Table 3). After PSM, all clinical variables, including age, became similar regardless of CH, and patients with CH still had poor OS compared to those without CH ( $p = 0.019$ ) (Figure 2C). Therefore, we can conclude that the presence of CH is an independent prognostic factor for OS in patients with advanced stage NSCLC.

In terms of the cause of death, we found that the significant difference according to the presence of CH was shown only in lung cancer mortality ( $p = 0.017$ ), especially in patients without evident cancer progression ( $p < 0.001$ ). This means that although recurrence had occurred, deaths not due to cancer progression or those of unknown origin were significantly more common in patients with CH. Judging the fact that the majority of patients completed the adjuvant therapy with the good compliance to postoperative surveillance, it is speculated that deaths from unknown origin were due to acute events, such as cardiopulmonary disease, sepsis, or stroke, rather than a relatively slow progression of cancer. Thus, we believe that these findings support our hypothesis that various adverse outcomes related to CH are amplified by CTx or RTx in patients with cancer with CH, which in turn affects survival.

Most of cytotoxic chemotherapeutics including platinum-based compounds such as cisplatin target DNA replication machinery. As conventional chemotherapies are designed to kill rapidly dividing cells, they cause critical DNA damage resulting in subsequent cell death.<sup>21,22</sup> However, mutations in genes related to cancer such as TP53, PPM1D, and CHEK2 impairs cell death process which should be normally activated upon DNA damage, leading to a hematopoietic stem cell survival advantage in the setting of cytotoxic drugs.<sup>23,24</sup> Thus, cancer-related therapies are thought to influence evolutionary trajectories of emerging CH clones. A recent study evaluated the clonal dynamics of CH in response to cancer therapy through sequential sampling and reported that cancer therapy preferentially selected for mutated clones in the DNA damage response (DDR) genes and that these clones had lower competitive fitness relative to non-DDR gene mutations in the absence of cytotoxic or radiation therapy.<sup>25</sup>

Recent studies focusing on CH mutations have demonstrated its impacts after the classification of CH based on their association with PD in hematological malignancy, CH-PD, or CH non-PD.<sup>10,25-27</sup> According to the Coombs et al.'s report, although CH itself did not show an independent prognostic

effect, CH-PD was associated with worse prognosis, regardless of age, sex, and smoking.<sup>10</sup> As with these findings, in our study, the prognostic effect remained significant in CH-PD ( $p = 0.038$ ) after the adjustment by age and sex, whereas it was marginally significant in CH itself ( $p = 0.055$ ) (Table 2). Moreover, unlike CH itself, the negative prognostic effect of CH-PD was gradually increased as the number of mutations increased, indicating that CH-PD is a more sensitive prognostic factor for OS than CH itself (Table 2). Notably, the rate of CH-PD (75.3%) in our study was higher than that reported in previous studies (52%–67%), presumably due to the higher proportion of current/previous smokers among the lung cancer patients.<sup>10,27</sup> We believe that this is the main reason why CH itself, as well as CH-PD, was found to be an independent prognostic factor for OS in our study (Figure 2C).

An important outstanding question is how should physicians manage lung cancer patients with CH mutations who are indicated for adjuvant therapy? First, it should be proceeded to refine the patient group where adjuvant therapy is beneficial to prognosis even at the risk of survival loss due to CH. Considering this, we performed subgroup analysis after stratifying our patients according to their overall stage (Supplementary Figure 2). Regardless of the overall stage, patients with CH mutations had worse survival outcomes than those without CH, but a significant difference was observed only in patients with stage IIB (5-year rate: 46.3% vs. 74.5%,  $p = 0.003$ ). Considering the survival benefit for adjuvant therapy ranging from 4% to 15%,<sup>28</sup> we recommend that adjuvant therapy be determined more carefully by a multidisciplinary approach in stage II patients with high-risk CH mutations. In addition, patients who have a high risk of developing adverse outcomes should be distinguished from those who do not. Although there is no clear definition for high-risk CH, the presence of significant blood count abnormalities, a single CH mutation at a high VAF ( $> 10\%$ ), multiple CH mutations, variants in TP53 and PPM1D, DNMT3A variants, and hotspot mutations of IDH1/2 are considered to put patients in the high-risk group.<sup>2,29</sup> Treatments that inhibit the progression of related adverse outcomes might be a good supplement for managing patients with lung cancer with CH mutations. For example, the anti-inflammatory agent canakinumab, a humanized monoclonal antibody against IL-1 $\beta$ , is reported to reduce cardiovascular events and the rates of lung cancer in high-risk patients with atherosclerotic disease.<sup>30</sup> In patients with stage III, where systemic therapy is essential for treatment, aggressive adaptation of molecularly targeted interventions can be an alternative for those with high-risk CH mutations. Finally, individualized follow-up duration and thorough monitoring for adverse outcomes of CH, such as cardiovascular disease or secondary malignancy, are required for patients with high-risk CH mutations.<sup>29</sup>

This study had notable limitations. First, selection bias is inherent in a retrospective study from a single institution; however, as the data in this study were gathered prospectively, we aimed to minimize this bias as much as possible. Second, the number of patients enrolled in this study was

relatively small, which may raise the possibility of selection bias. Indeed, some findings, which might have seemed different, were not significant. Thirdly, there are concerns that the results of CH mutation could have been confounded by circulatory tumor DNA (ctDNA). However, there is little possibility that CH, mutations of DNA from buffy coat, is affected by ctDNA from plasma. Following centrifugation, buffy coat may contain small or trace amounts of ctDNA, but the length of ctDNA is shorter than 200 base pair, being fragmented and washed away during library preparation for NGS.<sup>31</sup> Sequencing of paired tissue sample may help to interpret the results, but it was not feasible due to the problem of cancer sample acquisition. Finally, some patients died of unknown origin, which limited the accurate assessment of CH-related adverse outcomes.

## **Conclusion**

Preoperatively existing CH mutations have a significant clinical impact on patients with NSCLC who received surgery followed by adjuvant therapy, which decreases the survival outcome. Research efforts to validate our results are encouraged, and will help to reestablish our approach to managing clonal hematopoiesis in adjuvant therapy settings for NSCLC.

## **Acknowledgments**

The biospecimen and data used in this study were provided by Asan Bio-Resource Center, Korea Biobank Network [2020–10(207) 1st].

## **Support**

Supported by Genome Opinion Inc.

## **Disclosure of conflicts of interest**

### **S.K**

Employment: Genome Opinion Inc.

Stock and other ownership interests: Genome Opinion Inc.

### **Y.K**

Leadership: Genome Opinion Inc.

Stock and other ownership interests: Genome Opinion Inc.

## References

1. Genovese G, Kahler AK, Handsaker RE, et al: Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* 371:2477-87, 2014
2. Park SJ, Bejar R: Clonal hematopoiesis in cancer. *Exp Hematol* 83:105-112, 2020
3. Jaiswal S, Ebert BL: Clonal hematopoiesis in human aging and disease. *Science* 366, 2019
4. Jaiswal S, Fontanillas P, Flannick J, et al: Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* 371:2488-98, 2014
5. Welch JS, Ley TJ, Link DC, et al: The origin and evolution of mutations in acute myeloid leukemia. *Cell* 150:264-78, 2012
6. Coombs CC, Zehir A, Devlin SM, et al: Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* 21:374-382.e4, 2017
7. Gillis NK, Ball M, Zhang Q, et al: Clonal haemopoiesis and therapy-related myeloid malignancies in elderly patients: a proof-of-concept, case-control study. *Lancet Oncol* 18:112-121, 2017
8. DeZern AE, Malcovati L, Ebert BL: CHIP, CCUS, and Other Acronyms: Definition, Implications, and Impact on Practice. *Am Soc Clin Oncol Educ Book* 39:400-410, 2019
9. Bejar R: CHIP, ICUS, CCUS and other four-letter words. *Leukemia* 31:1869-1871, 2017
10. Coombs CC, Zehir A, Devlin SM, et al: Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* 21:374-382 e4, 2017
11. Bray F, Ferlay J, Soerjomataram I, et al: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68:394-424, 2018
12. O'Callaghan DS, O'Donnell D, O'Connell F, et al: The role of inflammation in the pathogenesis of non-small cell lung cancer. *J Thorac Oncol* 5:2024-36, 2010
13. Goldstraw P, Chansky K, Crowley J, et al: The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* 11:39-51, 2016
14. Yun JK, Bok JS, Lee GD, et al: Long-term outcomes of upfront surgery in patients with resectable pathological N2 non-small-cell lung cancer. *Eur J Cardiothorac Surg* 58:59-69, 2020
15. Detterbeck FC, Boffa DJ, Kim AW, et al: The Eighth Edition Lung Cancer Stage Classification. *Chest* 151:193-203, 2017
16. Takahashi K, Wang F, Kantarjian H, et al: Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. *Lancet Oncol* 18:100-111, 2017
17. Jaiswal S, Natarajan P, Silver AJ, et al: Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *New England Journal of Medicine* 377:111-121, 2017
18. Ho DE, Imai K, King G, et al: Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal* 15:199-236, 2017
19. Austin PC, Stuart EA: Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 34:3661-79, 2015

20. Busque L, Mio R, Mattioli J, et al: Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* 88:59-65, 1996
21. Lee JK, Choi YL, Kwon M, et al: Mechanisms and Consequences of Cancer Genome Instability: Lessons from Genome Sequencing Studies. *Annu Rev Pathol* 11:283-312, 2016
22. Ahmad S: Platinum-DNA interactions and subsequent cellular processes controlling sensitivity to anticancer platinum complexes. *Chem Biodivers* 7:543-66, 2010
23. Hsu JI, Dayaram T, Tovy A, et al: PPM1D Mutations Drive Clonal Hematopoiesis in Response to Cytotoxic Chemotherapy. *Cell Stem Cell* 23:700-713 e6, 2018
24. Bondar T, Medzhitov R: p53-mediated hematopoietic stem and progenitor cell competition. *Cell Stem Cell* 6:309-22, 2010
25. Bolton KL, Ptashkin RN, Gao T, et al: Cancer therapy shapes the fitness landscape of clonal hematopoiesis. *Nat Genet* 52:1219-1226, 2020
26. Bowman RL, Busque L, Levine RL: Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies. *Cell Stem Cell* 22:157-170, 2018
27. Bolton KL, Koh Y, Foote MB, et al: Clonal hematopoiesis is associated with risk of severe Covid-19. *medRxiv*, 2020
28. Pirker R: Adjuvant chemotherapy in patients with completely resected non-small cell lung cancer. *Transl Lung Cancer Res* 3:305-10, 2014
29. Bolton KL, Gillis NK, Coombs CC, et al: Managing Clonal Hematopoiesis in Patients With Solid Tumors. *J Clin Oncol* 37:7-11, 2019
30. Ridker PM, MacFadyen JG, Thuren T, et al: Effect of interleukin-1beta inhibition with canakinumab on incident lung cancer in patients with atherosclerosis: exploratory results from a randomised, double-blind, placebo-controlled trial. *Lancet* 390:1833-1842, 2017
31. Wan JCM, Massie C, Garcia-Corbacho J, et al: Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nature Reviews Cancer* 17:223-238, 2017

**Table 1. Baseline characteristics of enrolled patients according to the presence of clonal hematopoiesis**

| Variables                         | Total<br>(n = 341) | CH (-)<br>(n = 264) | CH (+)<br>(n = 77) | P-value |
|-----------------------------------|--------------------|---------------------|--------------------|---------|
| <b>Age (years)</b>                | 60.2 ± 8.6         | 58.9 ± 8.5          | 64.5 ± 7.8         | < 0.001 |
| <b>Sex (male)</b>                 | 221 (64.8)         | 161 (61.0)          | 60 (77.9)          | 0.009   |
| <b>History of smoking</b>         | 199 (58.4)         | 148 (56.1)          | 51 (66.2)          | 0.144   |
| <b>Charlson comorbidity score</b> |                    |                     |                    | 0.554   |
| 0                                 | 270 (79.2)         | 212 (80.3)          | 58 (75.3)          |         |
| 1                                 | 57 (16.7)          | 41 (15.5)           | 16 (20.8)          |         |
| ≥ 2                               | 14 (4.1)           | 11 (4.2)            | 3 (3.9)            |         |
| <b>Pulmonary function</b>         |                    |                     |                    |         |
| FEV1 < 60%                        | 15 (4.4)           | 10 (3.8)            | 5 (6.5)            | 0.486   |
| DLCO < 60%                        | 17 (5.2)           | 11 (4.3)            | 6 (8.1)            | 0.325   |
| <b>Histologic structure</b>       |                    |                     |                    | 0.851   |
| ADC*                              | 236 (69.2)         | 184 (69.7)          | 52 (67.5)          |         |
| SqCC*                             | 83 (24.3)          | 64 (24.2)           | 19 (24.7)          |         |
| Others                            | 22 (6.5)           | 16 (6.1)            | 6 (7.8)            |         |
| <b>Maximum tumor size (mm)</b>    | 40.0 ± 17.9        | 40.2 ± 16.9         | 39.4 ± 21.1        | 0.740   |
| <b>EGFR mutation</b>              |                    |                     |                    | 0.452   |
| Yes                               | 243 (71.3)         | 185 (70.1)          | 58 (75.3)          |         |
| No                                | 98 (28.7)          | 79 (29.9)           | 19 (24.7)          |         |
| <b>Pathological T factor</b>      |                    |                     |                    | 0.081   |
| T1                                | 72 (21.1)          | 48 (18.2)           | 24 (31.2)          |         |
| T2                                | 166 (48.7)         | 136 (51.5)          | 30 (39.0)          |         |
| T3                                | 69 (20.2)          | 54 (20.5)           | 15 (19.5)          |         |
| T4                                | 34 (10.0)          | 26 (9.8)            | 8 (10.4)           |         |
| <b>Pathological N factor</b>      |                    |                     |                    | 0.231   |
| N1                                | 151 (44.3)         | 122 (46.2)          | 29 (37.7)          |         |
| N2                                | 190 (55.7)         | 142 (53.8)          | 48 (62.3)          |         |
| <b>Pathological stage</b>         |                    |                     |                    | 0.367   |
| IIB                               | 101 (29.6)         | 83 (31.4)           | 18 (23.4)          |         |
| IIIA                              | 187 (54.8)         | 140 (53.0)          | 47 (61.0)          |         |
| IIIB                              | 53 (15.5)          | 41 (15.5)           | 12 (15.6)          |         |

| Type of adjuvant therapy |            |            |           | 0.298 |
|--------------------------|------------|------------|-----------|-------|
| CTx                      | 175 (51.3) | 140 (53.0) | 35 (45.5) |       |
| CRTx                     | 166 (48.7) | 124 (47.0) | 42 (54.5) |       |

Data are presented as no. (%) unless noted otherwise. CH: Clonal hematopoiesis, FEV1: Forced expiratory volume during the first second, DLCO: Diffusing capacity for carbon monoxide, ADC: Adenocarcinoma, SqCC: Squamous cell carcinoma, EGFR: Epidermal growth factor receptor, CRTx: Chemoradiotherapy, CTx: Chemotherapy.

**Table 2. Sensitivity test for clonal hematopoiesis, clonal hematopoiesis-potential driver, and clonal hematopoiesis non-potential driver**

|   | Number<br>(+) vs. (-) | Univariable analysis for OS |         | Multivariable analysis for OS<br>(adjusted by age and sex) |         |
|---|-----------------------|-----------------------------|---------|--|---------|
|   |                       | HR* (95% CI)                | P-value | HR (95% CI)  | P-value |
| <b>1) Analysis following the presence of mutation</b> |                       |                             |         |  |         |
| CH  | 77 vs. 264            | 1.67 (1.17–2.39)            | 0.005   | 1.45 (0.99–2.11)   | 0.055   |
| CH–PD   | 58 vs. 283            | 1.70 (1.22–2.40)            | 0.002   | 1.47 (1.02–2.09)   | 0.038   |
| CH non-PD   | 27 vs. 314            | 1.43 (0.87–2.34)            | 0.157   | 1.00 (0.58–1.72)   | 0.997   |
| <b>2) Analysis following the number of mutations</b>  |                       |                             |         |  |         |
| CH  |                       |                             |         |  |         |
| –1 vs. 0  | 63 vs. 264            | 1.56 (1.05–2.33)            | 0.027   | 1.37 (0.90–2.06)   | 0.139   |
| –2 vs. 0  | 11 vs. 264            | 1.68 (0.78–3.61)            | 0.184   | 1.47 (0.68–3.21)   | 0.329   |
| –3 vs. 0  | 3 vs. 264             | 6.27 (1.96–20.04)           | 0.002   | 4.93 (1.51–16.07)  | 0.008   |
| CH–PD   |                       |                             |         |  |         |
| –1 vs. 0  | 52 vs. 283            | 1.64 (1.09–2.46)            | 0.018   | 1.41 (0.97–2.09)   | 0.072   |
| –2 vs. 0  | 6 vs. 283             | 3.66 (1.48–9.03)            | 0.005   | 3.22 (1.28–8.10)   | 0.013   |
| CH non-PD   |                       |                             |         |  |         |
| –1 vs. 0  | 24 vs. 314            | 1.42 (0.86–2.35)            | 0.176   | 0.96 (0.55–1.69)   | 0.891   |
| –2 vs. 0  | 3 vs. 314             | 1.59 (0.22–11.43)           | 0.642   | 1.96 (0.27–14.29)  | 0.505   |

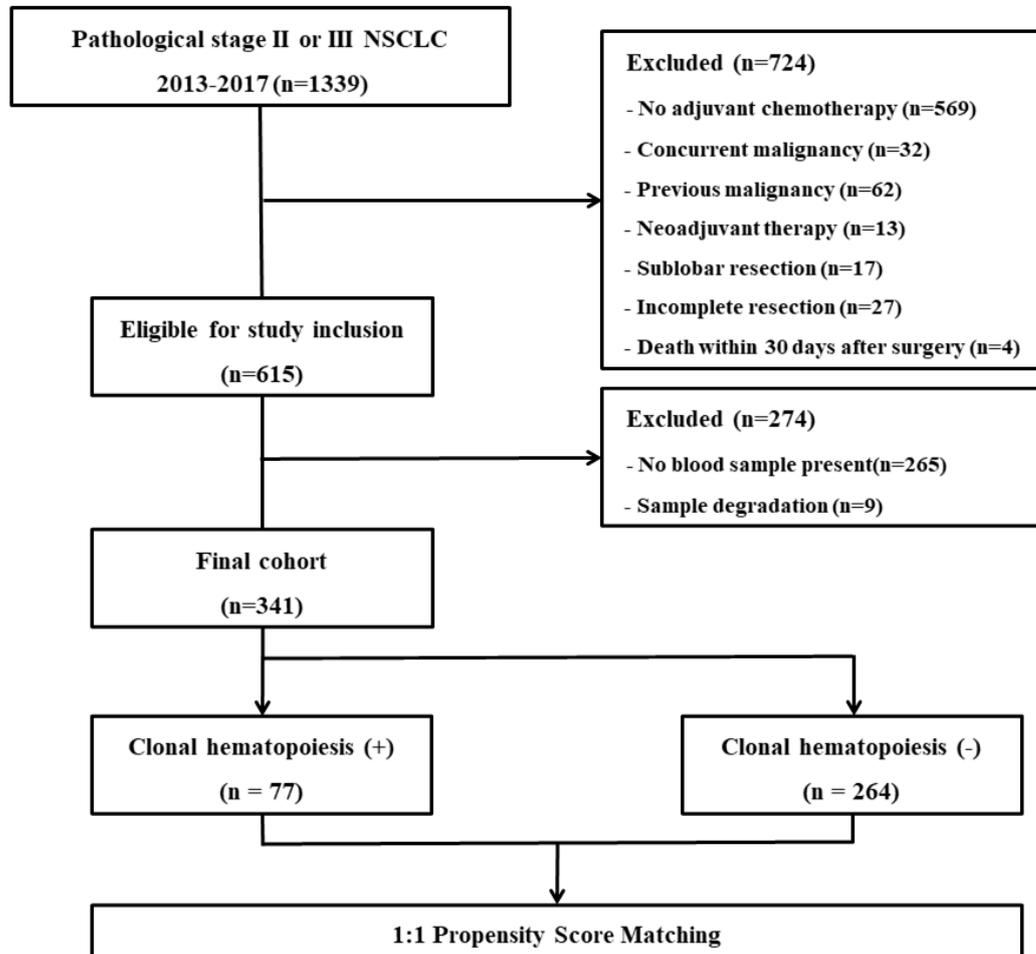
CH: Clonal hematopoiesis, PD: Potential driver, HR: Hazard ratio, CI: Confidence interval, VAF: Variant allele fraction

**Table 3. Univariable and multivariable analysis for overall survival of all patients**

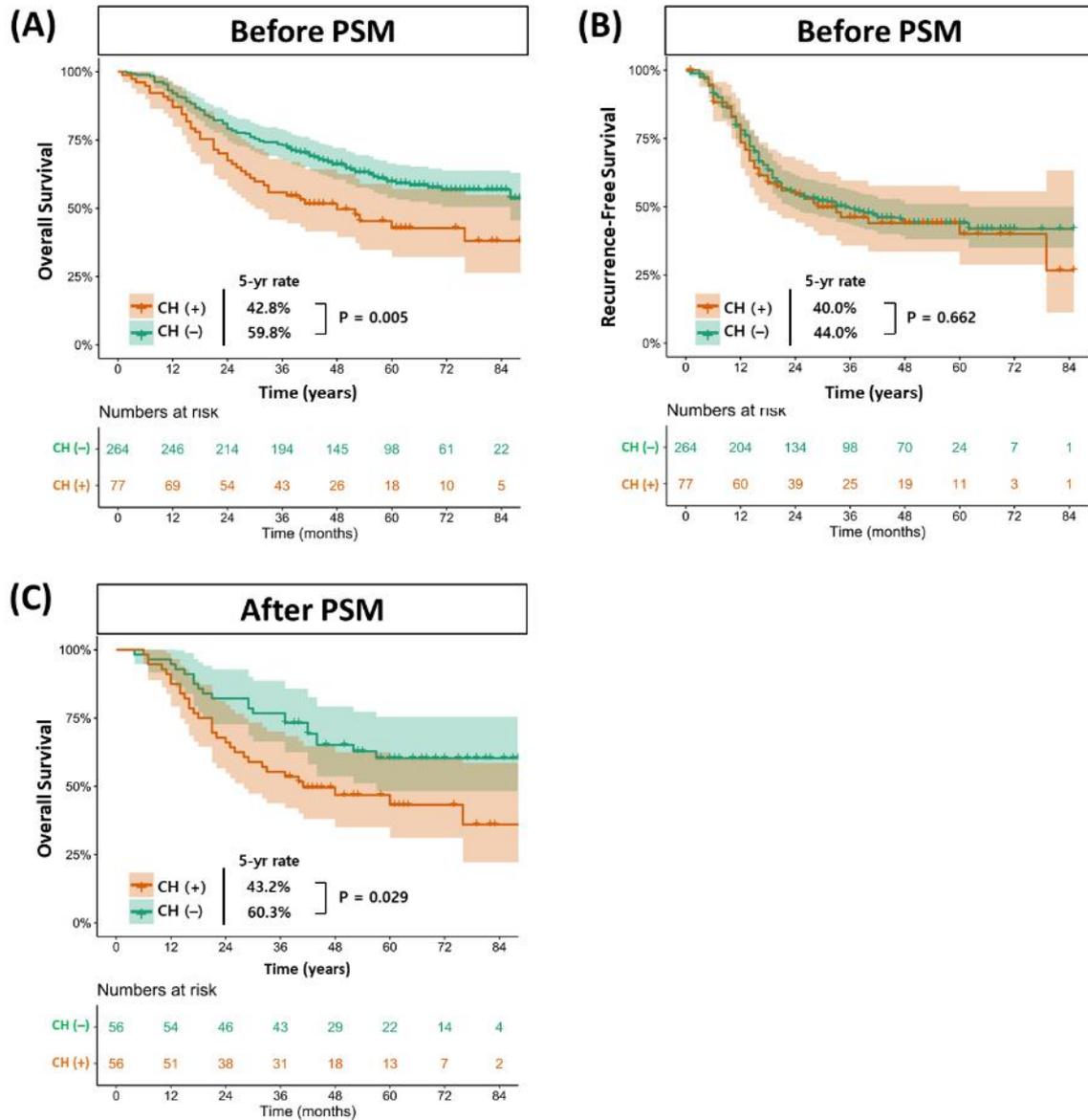
|                                   | Univariable analysis |         | Multivariable analysis |         |
|-----------------------------------|----------------------|---------|------------------------|---------|
|                                   | HR* (95% CI)         | P-value | HR (95% CI)            | P-value |
| <b>Presence of CH</b>             | 1.67 (1.17–2.39)     | 0.005   | 1.60 (1.12–2.30)       | 0.011   |
| <b>Age (years)</b>                | 1.03 (1.01–1.05)     | 0.001   |                        |         |
| <b>Sex (male)</b>                 | 1.10 (0.78–1.54)     | 0.595   |                        |         |
| <b>History of smoking</b>         | 1.20 (0.86–1.67)     | 0.290   |                        |         |
| <b>Charlson comorbidity score</b> |                      |         |                        |         |
| 1 vs. 0                           | 1.29 (0.85–1.95)     | 0.227   |                        |         |
| ≥ 2 vs. 0                         | 1.70 (1.14–2.53)     | 0.009   |                        |         |
| <b>Pulmonary function</b>         |                      |         |                        |         |
| FEV1 < 60%                        | 1.08 (0.51–2.32)     | 0.834   |                        |         |
| DLCO < 60%                        | 1.29 (0.85–1.95)     | 0.227   |                        |         |
| <b>EGFR mutation</b>              | 1.10 (0.76–1.59)     | 0.621   |                        |         |
| <b>Histologic structure</b>       |                      |         |                        |         |
| Sqcc vs. ADC                      | 0.88 (0.59–1.32)     | 0.532   | 0.97 (0.65–1.46)       | 0.889   |
| Others vs. ADC                    | 2.38 (1.36–4.18)     | 0.002   | 2.25 (1.28–3.95)       | 0.005   |
| <b>Pathological stage</b>         |                      |         |                        |         |
| IIIA vs. IIB                      | 1.93 (1.27–2.94)     | 0.002   | 1.79 (1.17–2.75)       | 0.007   |
| IIIB vs. IIB                      | 2.79 (1.68–4.63)     | < 0.001 | 2.60 (1.56–4.34)       | < 0.001 |
| <b>Adjuvant therapy</b>           |                      |         |                        |         |
| CRTx vs. CTx                      | 0.86 (0.62–1.19)     | 0.363   |                        |         |

OS: Overall survival, HR: Hazard ratio, CI: Confidence interval, CH: Clonal hematopoiesis, ADC: Adenocarcinoma, SqCC: Squamous cell carcinoma, CH: Clonal hematopoiesis, CRTx: Chemoradiotherapy, CTx: Chemotherapy.

Figure 1. CONSORT diagram.

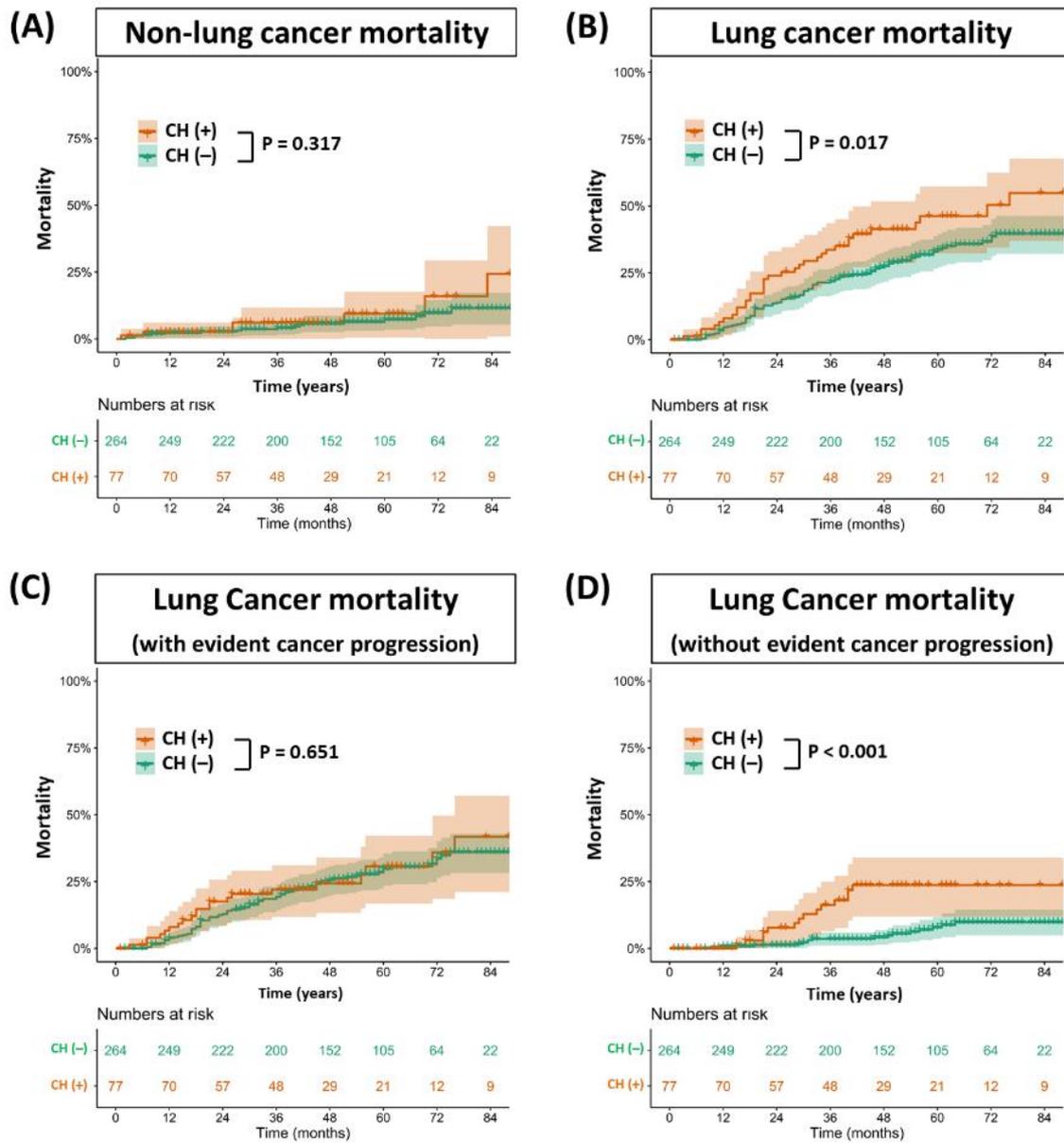


**Figure 2. Overall survival (A) and recurrence-free survival (B) of patients following the presence of CH mutations in the entire cohort. Overall survival (C) of patients following the presence of CH mutations after propensity score matching.**



CH: Clonal hematopoiesis.

**Figure 3. (A) Cumulative non-lung cancer mortality according to the presence of CH. (B) Cumulative lung cancer mortality according to the presence of CH. (C) Cumulative lung cancer mortality with evident cancer progression according to the presence of CH. (D) Cumulative lung cancer mortality without evident cancer progression according to the presence of CH.**



CH: Clonal hematopoiesis.

## **Supplementary Methods**

### **Next generation sequencing for CH**

The quantity of extracted DNA was assessed using a Qubit dsDNA BR Assay Kit (Invitrogen, Waltham, MA). A total of 150 ng input DNA per sample was fragmented by enzyme digestion using the SureSelect XT HS and XT Low Input Enzymatic Fragmentation Kit (Agilent, Santa Clara, CA). Following the SureSelect XT HS Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library protocol (Version C2, July 2019), the random fragmented samples were ligated to Illumina adaptors and PCR amplified using SureSelect XT HS Index Primers and Herculase II Fusion DNA polymerase. The cycling conditions were as follows: 98°C for 2 min; followed by 8 cycles of 98°C for 30 s, 60°C for 30 s, and 72°C for 1 min; and a final extension at 72°C for 5 min. Amplified-libraries were purified using AMPure XP beads (Beckman Coulter, Inc., Brea, CA). The quality and quantity of pre captured libraries were determined by TapeStation 4200 using a D1000 ScreenTape (Agilent, Santa Clara, CA). Next, pools of 16 indexed samples were concentrated to 3 µg in 12 µl DW using a Speedvac machine (Thermo Scientific, Waltham, MA) and hybridized with SureSelect XT2 Custom Capture library (0.5–2.9 Mb, Design ID: 3253531) for 16 h at 65°C. After hybridization, the captured targets were pulled down by Dynabeads MyOne Streptavidin T1 magnetic beads (ThermoFisher Scientific, Waltham, MA). The beads with captured DNA were washed once with wash buffer 1 and six times with wash buffer 2 to remove non-specific binding. The selected regions were amplified using SureSelect post capture primer mix and Herculase II Fusion DNA polymerase. The cycling conditions were as follows: 98°C for 2 min; followed by 12 cycles of 98°C for 30 s, 60°C for 30 s, and 72°C for 1 min; and a final extension at 72°C for 5 min. Libraries were identified with an Agilent TapeStation 4200 using High Sensitivity D 1000 ScreenTape (Agilent, Santa Clara, CA) and KAPA Library Quantification Kit (Kapa Biosystems).

### **Preprocessing, quality control analysis, and sample identification**

Targeted sequencing reads for PBMC samples were demultiplexed using Illumina's bcl2fastq (v2.17.1.14) to generate FASTQ files. We used SeqPrep for adapter trimming (default settings) and Sickle (v1.33) for low BQV base trimming (minimum average BQV = 30). Subsequently, trimmed FASTQ files were submitted to the GATK best practice pipeline, which includes alignment to the hg19 reference with BWA-MEM (v0.7.10). For all samples, duplicate marking and sorting was performed using PICARD (v1.94) MarkDuplicates, followed by indel realignment and base quality score recalibration using GATK Light (v2.3.9). Duplicate marking was performed again, resulting in a final coordinate sorted BAM per sample as an analysis ready BAM. Duplication metrics and BAM quality metrics were computed using PICARD (v1.94; MarkDuplicates, CalculateHsMetrics, CollectGcBiasMetrics). To specifically assess the potential sample labeling error (or sample

swapping), we applied an in-house script, which validated the genetic concordance in the sample VAF in a profile-wise manner. Samples that showed a high VAF correlation coefficient ( $> 0.90$ ) were considered as a sample (or samples) originated from the same patient.

### **Somatic mutation calling and filtering**

Analysis ready BAM files were processed through a somatic variant calling pipeline consisting of SNVer (0.4.1), LoFreq (0.6.1), and GATK UnifiedGenotyper (v2.3.9) for SNVs, and also an in-house InDel caller, which was used in previous cancer genome studies. To achieve comprehensive somatic variant calling, we enforced union between SNV callers. Technically, the requirements for positive SNVs/Indels were total reads  $\geq 10$ , Alt reads  $\geq 10$ , positive Alt reads  $\geq 5$ , negative Alt reads  $\geq 5$ , and VAF between 2% and 30%. We further filtered tri-allelic sites and common germline variants with  $MAF > 2\%$  in gnomAD, the 1000 Genomes Project release 3, ESP6500, and the ExAC (Exome Aggregation Consortium server). Finally, we filtered a subset of artifactual calls with  $MAF > 2\%$  in somatic variants on 1,000 CHIP negative cohort data, which is similar to Mutect2's normal option panel.

Analysis ready BAM files were processed through a somatic variant calling pipeline consisting of SNVer (0.4.1), LoFreq (0.6.1), and GATK UnifiedGenotyper (v2.3.9) for SNVs, and also an in-house InDel caller, which was used in our previously published cancer genome studies. To achieve comprehensive somatic variant calling, we enforced merged output between SNV callers.

Technically, the requirements for positive SNVs/Indels were total reads  $\geq 10$ , Alt reads  $\geq 10$ , positive Alt reads  $\geq 5$ , negative Alt reads  $\geq 5$ , MQV  $\geq 30$ , BQV  $\geq 30$ , and VAF between 2% and 30%. We further filtered tri-allelic sites and common germline variants with  $MAF > 2\%$  in gnomAD, the 1000 Genomes Project release 3, ESP6500, and the ExAC (Exome Aggregation Consortium server). Finally, we filtered a subset of artifactual calls with  $MAF > 2\%$  in somatic variants on 1,000 CHIP negative cohort data, which is similar to Mutect2's normal option panel.

**Supplementary Table 1. Eighty-nine clonal hematopoiesis genes targeted in the NGS panel.**

|        |         |          |        |         |        |          |        |
|--------|---------|----------|--------|---------|--------|----------|--------|
| APC    | ASXL1   | ASXL2    | ATM    | BCL11B  | BCOR   | BCORL1   | BIRC3  |
| BRAF   | BRCC3   | CARD11   | CASP8  | CBL     | CD58   | CD79B    | CNOT3  |
| CREBBP | CUX1    | DDX3X    | DNMT3A | EP300   | ETV6   | EZH2     | FAM46C |
| FBXW7  | FLT3    | FOXP1    | GNAS   | GNB1    | GPS2   | HIST1H1C | IDH2   |
| IKZF1  | IKZF2   | JAK1     | JAK2   | JAK3    | JARID2 | KDM6A    | KIT    |
| KLHL6  | KRAS    | LUC7L2   | MAP3K1 | KMT2D   | MPL    | MYD88    | NF1    |
| NFE2L2 | NOTCH1  | NOTCH2   | NRAS   | PDS5B   | PDSS2  | PHF6     | PHIP   |
| PIK3CA | PIK3R1  | PPM1D    | PRDM1  | PRPF40B | PTEN   | PTPN11   | RAD21  |
| RIT1   | RPS15   | SETD2    | SETDB1 | SF1     | SF3A1  | SF3B1    | SMC1A  |
| SMC3   | SRSF2   | STAG1    | STAG2  | STAT3   | SUZ12  | TBL1XR1  | TET1   |
| TET2   | TNFAIP3 | TNFRSF14 | TP53   | U2AF1   | VHL    | WT1      | ZRSR2  |
| CHEK2  |         |          |        |         |        |          |        |

NGS: next generation sequencing.

**Supplementary Table 2. Sensitivity test for clonal hematopoiesis based on the variant allele fraction value.**

|                 | Number      | Univariable analysis for OS |         | Multivariable analysis for OS<br>(adjusted by age and sex) |         |
|-----------------|-------------|-----------------------------|---------|--|---------|
|                 | (+) vs. (-) | HR* (95% CI)                | P-value | HR (95% CI)  | P-value |
| VAF $\geq$ 1%   | 134 vs. 207 | 1.21 (0.87–1.67)            | 0.261   | 1.05 (0.75–1.48)   | 0.773   |
| VAF $\geq$ 1.5% | 91 vs. 250  | 1.43 (1.01–2.03)            | 0.045   | 1.22 (0.84–1.77)   | 0.297   |
| VAF $\geq$ 1.8% | 80 vs. 261  | 1.71 (1.22–2.44)            | 0.003   | 1.49 (1.03–2.17)   | 0.035   |
| VAF $\geq$ 2%   | 77 vs. 264  | 1.67 (1.17–2.39)            | 0.005   | 1.45 (0.99–2.11)   | 0.055   |
| VAF $\geq$ 5%   | 33 vs. 324  | 1.77 (1.11–2.84)            | 0.018   | 1.56 (0.97–2.52)   | 0.069   |
| VAF $\geq$ 10%  | 17 vs. 324  | 1.47 (0.77–2.79)            | 0.243   | 1.38 (0.73–2.63)   | 0.324   |

CH: Clonal hematopoiesis, HR: Hazard ratio, CI: Confidence interval, VAF: variant allele fraction

**Supplementary Table 3. Baseline characteristics of enrolled patients according to the presence of clonal hematopoiesis after propensity score matching**

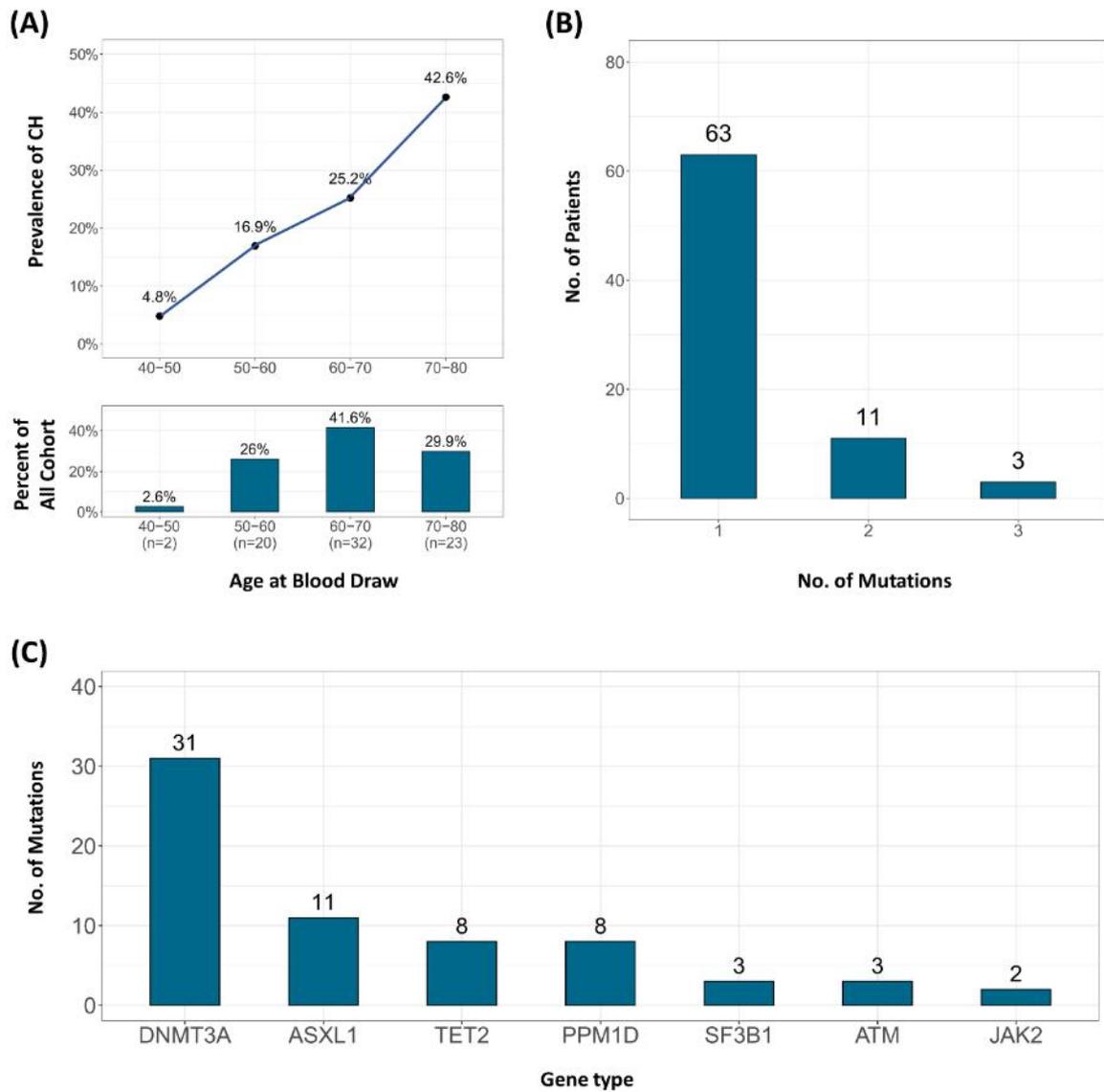
| Variables                         | CH (-)<br>(n = 56) | CH (+)<br>(n = 56) | P-value | ASMD    |
|-----------------------------------|--------------------|--------------------|---------|---------|
| <b>Age (years)</b>                | 62.0 ± 7.9         | 63.0 ± 8.1         | 0.481   | 0.134   |
| <b>Sex (male)</b>                 | 43 (76.8)          | 41 (73.2)          | 0.827   | 0.083   |
| <b>History of smoking</b>         | 36 (64.3)          | 37 (66.1)          | 1.000   | 0.037   |
| <b>Charlson comorbidity score</b> |                    |                    | 0.831   | 0.092   |
| 0                                 | 42 (75.0)          | 44 (78.6)          |         |         |
| 1                                 | 12 (21.4)          | 10 (17.9)          |         |         |
| ≥ 2                               | 2 (3.6)            | 2 (3.6)            |         |         |
| <b>Pulmonary function</b>         |                    |                    |         |         |
| FEV1 < 60%                        | 3 (5.4)            | 4 (7.1)            | 0.892   | 0.022   |
| DLCO < 60%                        | 4 (7.1)            | 4 (7.1)            | 1.000   | < 0.001 |
| <b>Histologic structure</b>       |                    |                    | 0.850   | 0.072   |
| ADC*                              | 43 (76.8)          | 38 (67.9)          |         |         |
| SqCC*                             | 10 (17.9)          | 14 (25.0)          |         |         |
| Others                            | 4 (7.1)            | 4 (7.1)            |         |         |
| <b>Maximum tumor size (mm)</b>    | 38.7 ± 16.6        | 39.6 ± 18.5        | 0.768   | 0.056   |
| <b>EGFR mutation</b>              |                    |                    | 1.000   | 0.039   |
| Yes                               | 39 (69.6)          | 40 (71.4)          |         |         |
| No                                | 17 (30.4)          | 16 (28.6)          |         |         |
| <b>Pathological stage</b>         |                    |                    | 0.827   | 0.117   |
| IIB                               | 16 (28.6)          | 15 (26.8)          |         |         |
| IIIA                              | 29 (51.8)          | 32 (57.1)          |         |         |
| IIIB                              | 11 (19.6)          | 9 (16.1)           |         |         |
| <b>Adjuvant chemotherapy</b>      |                    |                    | 0.855   | 0.067   |
| CTx                               | 32 (53.5)          | 30 (50.0)          |         |         |
| CRTx                              | 28 (46.7)          | 30 (50.0)          |         |         |

Data are presented as no. (%) unless noted otherwise. CH: Clonal hematopoiesis, ASMD: Absolute standardized mean difference, FEV1: Forced expiratory volume during the first second, DLCO: Diffusing capacity for carbon monoxide, ADC: Adenocarcinoma, SqCC: Squamous cell carcinoma, CRTx: Chemoradiotherapy, CTx: Chemotherapy.

**Supplementary Table 4. Cause of deaths for the entire observation period.**

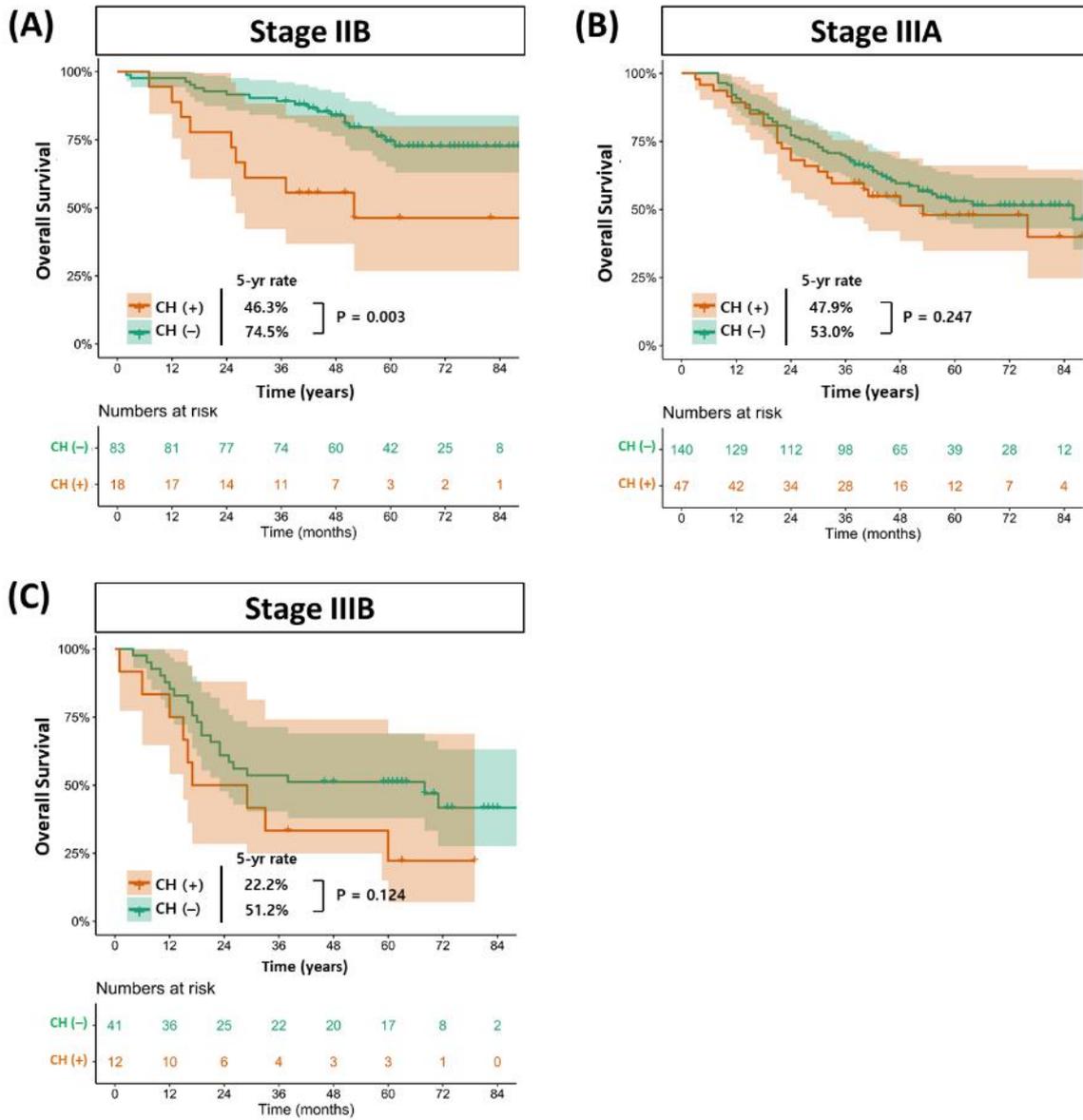
| Variables                      | Total<br>(n = 341) | CH (-)<br>(n = 264) | CH (+)<br>(n = 77) | P-value |
|--------------------------------|--------------------|---------------------|--------------------|---------|
| <b>Non-lung cancer related</b> | 25 (7.3)           | 18 (6.8)            | 7 (9.1)            | 0.671   |
| Cardiopulmonary                | 14 (4.1)           | 10 (3.8)            | 4 (5.2)            | 0.528   |
| Secondary malignancy           | 5 (1.5)            | 3 (1.1)             | 2 (2.6)            | 0.316   |
| Infection                      | 4 (1.2)            | 3 (1.1)             | 1 (1.3)            | 0.874   |
| Stroke                         | 2 (0.6)            | 1 (0.4)             | 1 (1.3)            | 0.401   |
| <b>Lung cancer related</b>     | 122 (35.8)         | 87 (33.0)           | 35 (45.5)          | 0.060   |
| Evident cancer progression     | 91 (26.7)          | 71 (26.9)           | 20 (26.0)          | 0.988   |
| Death after recurrence         | 7 (2.1)            | 2 (0.8)             | 5 (6.5)            | 0.008   |
| Unknown                        | 24 (7.0)           | 14 (5.3)            | 10 (13.0)          | 0.039   |

**Supplementary Figure 1. Characteristics of CH mutations identified in the overall cohort. (A) CH association with age, and age distribution of patients in the cohort. (B) Number of mutations harbored per patient. (C) The contribution of individual genes to the total number of potential driver somatic mutations that were observed.**



CH: Clonal hematopoiesis.

**Supplementary Figure 2. Overall survival of patients following the presence of CH mutations in (A) stage IIB, (B) stage IIIA, and (C) stage IIIB.**



CH: Clonal hematopoiesis.

## Korean abstracts (국문요약)

**서론:** 혈액 세포의 클론 개체수의 팽창에서 파생된 체세포 돌연변이인 클론성 조혈증(clonal hematopoiesis, 이하 CH)은 노화, 혈액 암 및 심혈관 질환 발생률 증가, 그리고 생존율 감소와 관련이 있다. 또한, CH는 고형 암 환자에서 더 흔하며, 약 30%는 혈액에 CH 돌연변이를 가지고 있다. 본 연구에서는 수술적 절제 후 보조 요법을 받은 비소세포 폐암 환자에서 수술 전 진단된 CH가 생존에 미치는 임상적 영향을 평가하였다.

**연구 방법:** 단일 기관 후향적 연구로 2013년부터 2017년까지 수술적 절제를 시행하고 보조치료를 받은 비소세포성 폐암 환자 341 명의 혈액 샘플과 의료기록을 후향적으로 검토하고 이를 분석했다. 보조 항암치료는 백금 기반의 화학요법으로 수술 후 4-6 주 동안 총 4 주기의 치료로 권장되었다. CH 돌연변이의 임상적 관련성을 평가하기 위해 수술 전에 수집된 341 명의 혈액 샘플에서 차세대 염기서열 분석을 통해 CH와 연관된 것으로 밝혀진 89 개의 유전자를 분석하였다. CH 존재 여부로 분류된 두 그룹 간에 가능한 선택 편향을 최소화하기 위해 프로펜시티 점수 매칭 기법을 채택했다.

**결과:** 비소세포성 폐암 환자의 23% (77/341)에서 수술 전 CH가 존재하는 것으로 확인되었다. 병리학적 병기는 IIB가 101명 (29.6%), IIIA가 187명 (54.8%), 그리고 IIIB가 53명 (15.5%)이었다. CH 돌연변이가 진단된 환자는 그렇지 않은 환자들에 비해 전체 생존율이 유의하게 낮았으며 (5년 생존율: 42.8% vs. 59.8%,  $p = 0.019$ ) 이는 프로펜시티 점수 매칭으로 보정한 후에도 비슷한 결과를 보였다 (5년 생존율: 43.2% vs. 58.7%,  $p = 0.019$ ). 사망 원인에 따르면, CH가 있는 경우 암 관련 사망률의 증가와 관련이 있었으며 ( $p = 0.017$ ), 그 중에서도 암 진행이 명확치 않은 암 관련 사망률과 관련이 있었다 ( $p < 0.001$ ). 다 변량 분석에서는 조직학적 아형과 종양의 병기와 더불어 CH 여부가 수술 후 보조 항암치료를 받은 비소세포성 폐암 환자들에 있어 유의한 예후인자인 것으로 확인되었다 (위험비 [95% 신뢰구간] = 1.60[1.12-2.30],  $p = 0.011$ ). 연령은 단 변량 분석에서 유의한 결과를 보였으나 다 변량 분석에서는 CH 여부를 비롯해 다른 공변량을 보정하자 유의하지 않게 되었다.

**결론:** 수술적 절제 후 보조 요법을 시행한 비소세포성 폐암 환자에서, 수술 전 존재하는 CH는 불량한 생존 결과와 유의한 관계가 있다.