RESEARCH ARTICLE

**Editorial Process:** Submission:09/12/2024 Acceptance:06/03/2025

# High Proportion of Cytomegalovirus DNA from Tissue Samples of Non Small Cell Lung Carcinoma in Persahabatan Hospital National Respiratory Center, Indonesia

**Ibrahim Nur Insan Putra Dharmawan[1], Jamal Zaini[1]*, Irandi Putra Pratomo[1], Eylin Rahardjo[1], Hana Khairina Putri Faisal[1], Dicky Soehardiman[1], Erlang Samoedro[1], Triya Damayanti[1], Diah Handayani[1], Fanny Fachrucha[1], Dimas Ramadhian Noor[2], Sita Laksmi Andarini[1]**

## Abstract

**Background:** Cytomegalovirus (CMV) is a virus with high seroprevalence in general population. Recent evidence shows that CMV is linked to various types of cancer, including lung cancer. This study aims to determine the relationship between CMV and non-small cell lung carcinoma (NSCLC). Data from this study will be useful for further research in elucidating the link between CMV and lung cancer. **Method:** This research was an observational study using a cross-sectional method to determine the proportion of CMV DNA in NSCLC tissue samples. Formalin-fixed paraffin-embedded (FFPE) tissue samples were taken from archival at Persahabatan Hospital on 2017-2023. The detection of CMV was carried out using polymerase chain reaction (PCR) and electrophoresis. Accompanying data was taken from medical records. **Results:** A total of 87 tissue samples from 87 different subjects were included in this study. Most of the research subjects were male smokers, had a heavy Brinkman index with an average age of 59.1 years. The proportion of CMV DNA detected in FFPE samples was 21%. The proportion of CMV DNA was higher in tissue samples with positive *EGFR* mutations although not statistically significant. The proportion of CMV DNA was not related to smoking status, Brinkman index, tissue sampling method, and NSCLC subtype. **Conclusion:** High proportion of cytomegalovirus DNA were detected in NSCLC FFPE samples in Indonesia. The proportion of CMV DNA was higher in NSCLC with *EGFR* mutations.

**Keywords:** cytomegalovirus- epidermal growth factor receptor- non-small cell lung cancer

*Asian Pac J Cancer Prev,* **26 (6)**, 1987-1995

## Introduction

Cytomegalovirus (CMV) is a virus of the herpesviridae family that can infect humans and other primates. Human cytomegalovirus (HCMV) infection, also known as human herpesvirus 5, often becomes latent and lasts indefinitely [1]. Cytomegalovirus is a very common virus. Studies show that its seroprevalence among the general population ranges between 24.6 – 95.7 % among countries in South East Asia. Its seroprevalence is higher in older populations [2]. Viruses from the herpesvirus family have two phases of infection,the productive/lytic and the latent phase. In the productive phase, cells infected by CMV will be actively hijacked to produce more virions. In the latent phase, the CMV genetic material will be integrated with the cell's deoxyribonucleic acid (DNA). These two forms are influenced by the immediate early (IE) gene, a promoter called Major Immediate Early Promoter (*MIEP*) and a protein called pp71. The pp71 protein contained in the outer layer of CMV will enter the cell during the infection process. The pp71 protein will activate *MIEP* which will then activate the IE gene to produce IE1 and IE2 proteins. These two proteins will push CMV infection tendency to a productive phase. Subsequently, the virus can invade other cells with the help of glycoprotein B produced by the *gB* gene located on the viral envelope. The glycoprotein B binds to the epidermal growth factor receptor (*EGFR*) as a mechanism of viral entry [3].

Oncogenic viruses are an important etiology of malignancy, including lung cancer. Oncogenic viruses cause latent infections in human cells and gradually cause modifications to the genetic structure of cells,

*[1]Department of Pulmonology and Respiratory Medicine Universitas Indonesia, Persahabatan Hospital National Respiratory Center, Indonesia. [2]Human Cancer Research Center, Indonesia Medical Education and Research Institute, Indonesia. *For Correspondence: jamal.zaini@gmail.com*

immune responses and cell death processes which result in changes in the nature of cells to become malignant. Currently there are seven recognized oncogenic viruses, which are Epstein–Barr virus (EBV), hepatitis B virus (HBV), human T-lymphotropic virus 1 (HTLV-1), human papillomaviruses (HPVs), hepatitis C virus (HCV), Kaposi sarcoma-associated herpesvirus (KSHV) dan Merkel cell polyomavirus (MCPyV) [4]. Cytomegalovirus is not currently recognized as an oncogenic virus. However, its genetic material can be found in various malignancies such as glioblastoma, breast, colon, and prostate cancer. In these studies, CMV DNA was found in the affected tissue and not found in the surrounding healthy cells. These findings suggest that CMV might plays a role in carcinogenesis. Research studying valganciclovir, an antiviral therapy, as adjuvant therapy for glioblastoma patients with detected CMV DNA also shows promising results with improved outcomes in these patients [5, 6].

Evidences examining the relationship between CMV and malignancies in the airways are scarce. In the airway system, CMV deoxyribonucleic acid (DNA) material was found in 35.3% of the histopathological specimens of nasopharyngeal cancer studied [7]. In lung cancer patients after tumor resection, traces of CMV infection in resected specimens were associated with an increased risk of more rapid tumor recurrence. One study examining the proportion of detected CMV DNA in lung adenocarcinoma cytology samples shows a proportion of 20.8% [8]. More evidence regarding CMV in lung cancer is needed. This study aims to determine the relationship between CMV and non-small cell carcinoma lung cancer in Indonesia. We hope that the data from this study will be useful for further research into the link between CMV and lung cancer in the future.

## Materials and Methods

*Subjects and samples*

This study was an observational study using a cross-sectional method to examine the proportion of CMV DNA in non-small cell lung carcinoma. The research was conducted using histopathological tissue samples stored as formalin-fixed paraffin-embedded (FFPE) in the Anatomic Pathology Laboratory of Persahabatan Hospital, Jakarta, Indonesia. The samples included in the study were all non-small cell lung carcinoma samples obtained in 2017-2023 through open surgery or tissue biopsy. The DNA from FFPE samples was extracted and underwent cytomegalovirus DNA detection at the Human Cancer Research Center at the Indonesian Medical Education and Research Institute (IMERI). Samples were excluded if the samples did not fulfill the minimum number of cells considered enough for DNA CMV detection according to the Pathology Anatomy expert, which is a minimum of 200 cells. Samples were also excluded if the final pathological diagnosis was inconclusive and still differentially diagnosed with small cell lung carcinoma. Accompanying data were taken from medical records.

*Extraction of Cells' DNA*

Cytomegalovirus DNA detection was carried out through several steps, DNA extraction, polymerase chain reaction (PCR) and gel electrophoresis. The DNA extraction was carried out using Quick-DNA™ FFPE kit (Zymo Research, California, USA) and done according to product protocols. At this stage, three stages were carried out, which were deparaffinization, tissue digestion, and DNA purification. First, the deparaffinization process was carried out to remove the paraffin substance that fixes the tissue using the deparaffinization solution provided in the kit. The sample was then lysed using proteinase K to separate DNA material from the tissue. The fixing was carried out at a temperature of 55°C for 4 hours to reverse the formalin crosslinking that occurred previously in the tissue fixation process. At the DNA purification stage, genomic lysis buffer was used. The sample was then centrifuged at a speed of 10,000g for 1 minute. Then the samples were washed with Genomic DNA Wash 1 and 2 solutions and centrifuged again with strengths of 10,000g and 12,000g, respectively. At this stage, pure DNA was collected. In this study, the amount of DNA obtained varied between 7 - 583 nanograms per microliter for each subject.

*Polymerase Chain Reaction*

The process continues with polymerase chain reaction (PCR) to increase the amount of DNA available for further processing. Primers were used to cut DNA and reproduce the desired segment. Amplicons from these slices were compared with positive controls.

The PCR were carried out with a ProFlex™ PCR System (Thermofisher Scientific, Massachusetts, United States). The DNA primers used were DNA primers from the major immediate early (*MIE*) and *gB* genes of the CMV virus (Human herpesvirus 5 strain Merlin, NCBI Reference Sequence: NC_006273.2). The *MIE* gene is located in the nucleotide sequence 170,689 to 174,090 in the CMV genome. The primers used to detect the *MIE* gene were the *MIE6* and *MIE7* primers which will produce a fragment of the *MIE* gene with a size of 110 base pairs (bp) long located in the nucleotide sequence 172,597 to 172,706. To detect the *gB* gene, *CMV 5A* and 5B primers were used which will produce a DNA fragment 326 bp long and located in the base sequence 82,641 to 82,966. The PCR kit used is Toyobo KOD One™ PCR Master Mix (Toyobo, Osaka, Japan). The PCR primer used in this study can be seen in Table 1. The primer used has been confirmed using Basic Local Alignment Search Tool (BLAST) from NCBI and no hairpin detected. The PCR condition we used were denaturation at 94°C, annealing at 60oC, and extension at 68°C. In this study, 40 PCR cycles were carried out for each sample.

*Detection of CMV DNA*

The amplified DNA samples were then tested by electrophoresis. This study used geneblock as a positive control. Geneblock is synthetic DNA made by synthetically building the *MIE* and *gB* genes interspersed with random nucleotide sequences between them. The Geneblock in this research was purchased through PT Genetika Science, Banten, Indonesia and manufactured by Integrated DNA Technologies, Iowa, United States.

The geneblock nucleotide arrangement used can be seen in Figure 1.

*Ethics*

This study has been approved by the Ethical Committee for Health Research of Persahabatan Hospital National Respiratory Center, Indonesia. Approval letter number 129/KEPK-RSUPP/09/2023. All subject's identities and data were kept confidential.

## Results

*Subjects characteristics*

Data collection, DNA extraction, CMV detection, and data analysis were done between October 2023 and January 2024 after the study was approved by the ethic committee. There were 16087 FFPE samples collected and stored in Pathology Anatomy Laboratorium between January 2017 and October 2023. After matching with inclusion and exclusion criteria, a total of 87 subjects were chosen in the study (Figure 2).

In this study, tissue samples were obtained using open surgical techniques, i.e, thoracotomy, or with minimally invasive biopsy techniques such as bronchoscopy. The proportion of the two techniques used was similar, 44 samples (50.5%) vs 43 samples (49.5%) for open surgical vs minimal invasive biopsy respectively.

The history of smoking was obtained from medical records. Smoking status was assigned by the clinician managing the patient and was defined as a history of smoking more than one cigarette per day for at least one year. As much as 70.5% of subjects had history of smoking. Most of the subjects were heavy smokers according to the Brinkman index. Data was also obtained regarding *EGFR* mutations. Of the 87 research subjects, 38 subjects were tested for *EGFR* mutations. More than half were detected as having positive *EGFR* mutations with 55% being common mutations (exon 19 deletion mutation and substitution mutation in exon 21 L858R), 3% uncommon mutations and the rest wild type. Detailed characteristics of the study subjects can be seen in Table 2.

*CMV DNA proportion*

The proportion of CMV was examined in this study as the main outcome. In this study, CMV DNA was detected in 18 samples out of a total of 87 samples studied (21%). Two genes were tested, which were the *MIE* and *gB* genes. A sample is said to be positive for CMV if there is a positive result for at least 1 of the 2 genes. Of the samples that were positive for CMV, 61% had the *gB* gene, 28% had the *MIE* gene and 11% had the *gB* and *MIE* genes (Figure 3).

The proportion of CMV in the sample was analyzed and related to several variables. Among samples with positive *EGFR* mutation results, the proportion of CMV was more than twice as high compared to those without *EGFR* mutations. However, this difference was not statistically significant (p = 0.43). In smokers and non-smokers, the proportion of CMV was almost similar. In different Brinkman index classifications, the proportion of CMV was similar. The proportion of CMV is higher in adenocarcinoma compared to squamous cell carcinoma and other types but this is not statistically significant. The proportion of CMV DNA found was also greater in newer samples compared to older samples. In this study, the subjects were divided into subjects who underwent tissue sampling within the last three years and before the last three years and found that the CMV DNA positivity rate was significantly higher in tissue samples taken within the last three years. (p=0.009) There was no significant

Table 1. Primer PCR Used to Detect MIE (MIE 6 and MIE 7) and gB (CMV 5A and CMV 5B) Gene and Its Product Size

| No | CMV gene primer | Nucleotide sequence 5'-3' | Product size (bp) |
|---|---|---|---|
| 1 | *MIE6* (F) | AGTGTGGATGACCTACGGGCCATCG | 110 |
| 2 | *MIE7* (R) | GGTGACACCAGAGAATCAGAGGAGC | |
| 3 | *CMV 5A* (F) | TCATGAGGTCGTCCAGA | 326 |
| 4 | *CMV 5B* (R) | TGAGGAATGTCAGCTTC | |

GGTCCTTGAAATTCGGCTCA**GGTGACACCAGAGAATCAGAGGAGC**TGGCACCAGCGGTGGCCAAA
GTGTAGGCTACAATAGCCTCTTCCTCATCTGACTCCTCGG**CGATGGCCCGTAGGTCATCCACACT**AG
GAGAGCAGATGGCGCAGAACGGCAATGAGTGTTTGACAC**TCATGAGGTCGTCCAGA**CCCTTGAGG
TAGGGCGGTAGCGGGTCGACTACCTTGTCCTCCACGTACTTTACCCGCTGCTTGTACGAGTTGAATTC
GCGCATGATCTCTTCGAGGTCAAAAACGTTGCTGGAACGCAGCTCTTTCTGCGAGTAAAGTTCCAGT
ACCCTGAAGTCGGTATTTTCCAGCGGGTCGATATCCAGGGCGATCATGCTGTCGACGGTGGAGATA
CTGCTGAGGTCAATCATGGTTTGAAGAGGTAGTCCACGTACTCGTAGGCCGAGTTCCCGGCGATGA
AGATCTTGAGGCTGG**GAAGCTGACATTCCTCA**GTTAGTTCGAGCGTAATGTCACAAA

Figure 1. Synthetic DNA (Geneblock) was Used as a Positive Control. The base sequence marked with a continuous-line rectangle is the MIE gene primer. The dotted-line rectangle is the gB gene primer. The arrangement of bases outside the segment flanking the primer is a random arrangement of nucleotides as "filler"

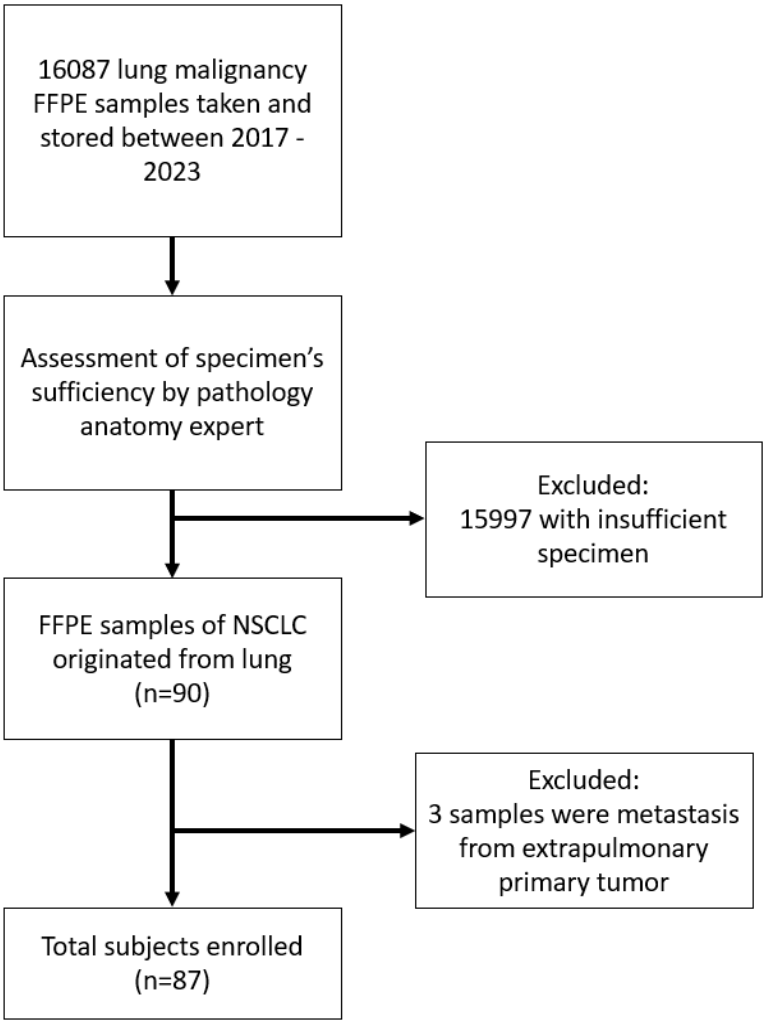


Figure 2. Sample Recruitment Workflow. Total of 87 Subjects were Enrolled in this Study

difference between the proportion of CMV DNA taken using surgical methods and minimally invasive biopsy. The details of CMV DNA proportion and its relation to the variable can be seen in Table 3.

## Discussion

In this study, CMV DNA detection was carried out in FFPE NSCLC samples to determine the proportion of CMV DNA in NSCLC. More evidences are mounting regarding the relationship between cytomegalovirus and malignancy in other types of cancer, for example, glioblastoma and nasopharyngeal cancer, but very little has been studied in lung cancer. We hope that the results of this study will be useful as a basis for further research in the future regarding the relationship between CMV

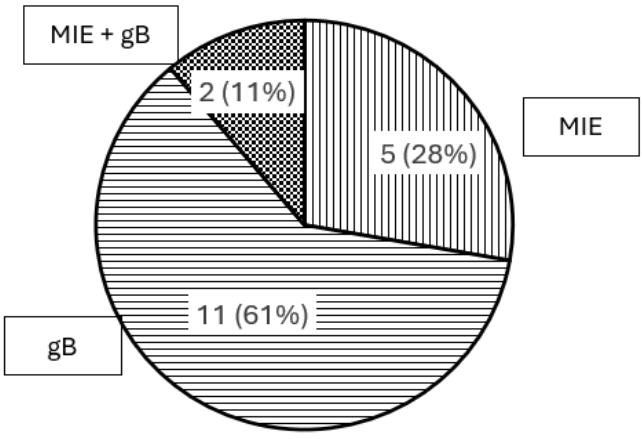


Figure 3. Type of CMV DNA Detected in FFPE Samples (N=18)

Table 2. Characteristics of Study Subjects

| Variable | N (87) | % |
|---|---|---|
| Sex | | |
| Male | 69 | 79.3 |
| Female | 18 | 20.7 |
| Age, years | | |
| Mean ± SD | 59.1 ± 10.2 | |
| Specimen acquisition technique (N=87) | | |
| Open surgery | 44 | 50.5 |
| Minimally-invasive biopsy | 43 | 49.5 |
| Transbronchial biopsy | 3 | 7.0 |
| Pleural biopsy | 4 | 9.3 |
| Bronchoscopy – forceps biopsy | 13 | 30.2 |
| Transthoracal core biopsy | 16 | 37.2 |
| Bronchoscopy – cryobiopsy | 7 | 16.3 |
| Year sample obtained (N=87) | | |
| 2017 | 7 | 8.0 |
| 2018 | 13 | 14.9 |
| 2019 | 7 | 8.0 |
| 2020 | 1 | 1.1 |
| 2021 | 5 | 5.7 |
| 2022 | 31 | 35.6 |
| 2023 | 23 | 26.4 |
| Smoking status (N=87) | | |
| Smoker | 43 | 49.0 |
| Non-smoker | 18 | 21.0 |
| Unknown | 26 | 30.0 |
| Brinkman Indeks (N=43) | | |
| Light | 4 | 9.3 |
| Moderate | 18 | 41.9 |
| Heavy | 20 | 46.5 |
| Unknown | 1 | 2.3 |
| EGFR mutation (N=38) | | |
| Wild type | 16 | 42.0 |
| Common mutation | 21 | 55.0 |
| Uncommon | 1 | 3.0 |
| NSCLC type (N=87) | | |
| Adenocarcinoma | 55 | 63.0 |
| Squamous cell carcinoma | 29 | 33.0 |
| Other | 3 | 4.0 |
| Stage at diagnosis (N=70) | | |
| IA | 3 | 4.3 |
| IB | 3 | 4.3 |
| IIA | 2 | 2.9 |
| IIB | 3 | 4.3 |
| IIIA | 9 | 12.8 |
| IIIB | 5 | 7.1 |
| IIIC | 2 | 2.9 |
| IVA | 37 | 52.9 |
| IVB | 6 | 8.6 |
| Performance status at diagnosis (N=87) | | |
| 0 | 9 | 10.3 |
| 1 | 41 | 47.1 |
| 2 | 11 | 12.6 |
| 3 | 8 | 9.2 |
| 4 | 1 | 1.2 |
| Unknown | 17 | 19.5 |

*Brinkman Index light: 0-200, moderate: 201-599, heavy: ≥600; EGFR: epidermal growth factor receptor; NSCLC: non small cell lung carcinoma.

and lung cancer, its management and its relationship to prognosis. In this research, the PCR and FFPE samples were chosen as samples because these can be obtained easily at relatively affordable costs.

This study included a total of 87 FFPE samples originating from the Anatomical Pathology laboratory at Persahabatan Hospital's archival. Tissue samples were obtained between 2017 and 2023. Samples were selected based on the total sampling method. The samples were selected based on expertise by pathologist and have sufficient specimen for DNA extraction with criteria of a minimum of 200 cells. Accompanying data, including demographic data, smoking history, cancer stage at diagnosis, *EGFR* mutation and NSCLC type, were obtained through hardcopy and electronic medical records. Of the 87 FFPE samples included in this study,

Table 3. Proportion of CMV DNA within Different Variables

| Variable | CMV | | P value |
|---|---|---|---|
| | Positive | Negative | |
| EGFR mutation (N=38) | | | |
| Positive | 6 | 16 | 0.43 |
| Negative | 2 | 14 | |
| Smoking status (N=61) | | | |
| Smoker | 11 | 32 | 1 |
| Nonsmoker | 4 | 14 | |
| Brinkman Indeks (N=59) | | | |
| Nonsmoker | 4 | 13 | 0.95 |
| Light | 1 | 3 | |
| Moderate | 4 | 14 | |
| Heavy | 6 | 14 | |
| NSCLC type (N=87) | | | |
| Adenocarcinoma | 13 | 42 | 0.53 |
| Squamous cell carcinoma | 5 | 24 | |
| Other | 0 | 3 | |
| Biopsy technique (N=87) | | | |
| Open surgery | 6 | 38 | 0.119 |
| Minimally invasive biopsy | 12 | 31 | |
| Time of sample biopsy (N=87) | | | |
| Within 3 years | 17 | 42 | 0.009 |
| More than 3 years | 1 | 27 | |

there were 61 samples that had complete accompanying data. However, all CMV DNA detection results from all samples were still reported as the main outcome considering the small number of samples obtained.

Most of study subjects was male (79.3%) with an average age of 59.1 years. The proportion of men in this study was higher than women. This is similar to various other studies which state that men are more susceptible to lung cancer than women, especially those aged over 54 years. 39 This is due to the higher level of tobacco consumption among men compared to women. Cigarettes are the main risk factor for lung cancer [9]. These results are in accordance with recommendations from the American Thoracic Society which recommends that adult patients aged 50 – 80 years who have a 20-pack-year smoking history undergo lung cancer screening [10].

The majority of subjects smoked with a heavy Brinkman index, in accordance with various literature which has stated that smoking is the main risk factor for lung cancer [11,12]. Bivariate analysis was conducted between CMV and smoking status and the Brinkman index but these were not statistically significant. These results are similar to research by Harabajsa, et al which found that there was no significant relationship between CMV and smoking in patients with lung cancer. This result is different from research conducted by Dickerson, et al which found that the smoking population had a higher proportion of CMV seroprevalence. This difference can be caused by differences in the examination methods. In a study by Dickerson, et al, the sample used was blood to detect antibodies to CMV while in our study, CMV DNA was directly detected in lung cancer tissue using PCR and electrophoresis methods [8, 13].

Tissue samples in this study were obtained through open-surgical and minimally invasive biopsy techniques in similar proportions. The results of this study show that CMV DNA can be found in tissue samples obtained by either of these methods. There was no significant difference regarding the CMV DNA positivity rate between the two tissue collection techniques. However, the CMV DNA positivity rate was higher in samples taken by minimally invasive biopsy methods. This could be because these samples were obtained within the last three years, while samples originating from surgical tissue were partially obtained in 2017-2021. This also shows that tissue samples taken with minimally invasive biopsy techniques can be used in CMV DNA testing.

Most of the subjects in this study were diagnosed at an advanced stage. This result concurs with data from guidelines by the Indonesian Lung Doctors Association (PDPI) which states that the majority of patients with lung cancer in Indonesia are diagnosed at an advanced stage [14]. The most common types of NSCLC in this study were adenocarcinoma and squamous cell carcinoma (SCC).

The proportion of *EGFR* mutations in this study was found to be 58% of all samples tested for *EGFR* mutations. This figure is greater than the figure compiled from a meta-analysis by Zhang, et al in 2016 which shows that the proportion of adenocarcinomas with *EGFR* mutations was found to be 38% in China. It is also higher than the proportion found in cytological sample of lung cancer in Indonesia (44.4%) and in Asia-Pacific (49.3%) [15–17]. In this study, it was also found that 95% of the *EGFR* mutations detected were exon 19 deletion and L858R substitution mutation in exon 21 (common mutations). These results are in accordance with various other studies which state that these two mutations are the most common types of *EGFR* mutations found [18].

The main outcome of this research is the proportion of CMV DNA found in the NSCLC samples. In this study, CMV DNA was detected in 21% of samples. The proportion of CMV is greater in adenocarcinoma than in SCC. These results are similar with Harabajsa et al who used the same PCR primers to detect CMV DNA. However, there are differences in the research methodology used. In the research conducted by Harabajsa, et al, the samples were cytology slides so that the sample preparation process did not include a formalin and deparaffinization process. In our study, the samples were FFPE samples that,during the manufacturing process, preserved using formalin at room temperature and embedded in a paraffin block. Nevertheless, the proportion of CMV DNA are similar in both studies which and either cytology or histology specimens may be used in future research [8].

Research conducted by Berrino, et al in 2020 suggested that the tissue preservation process using formalin at room temperature can disrupt the integrity of nucleic acids in tissue samples and their quality continues to decline over time. This research also shows that within six months of the tissue being taken, the integrity of nucleic acids in formalin-fixed tissue samples progressively decreases, especially for detecting longer DNA segments. Another study by Watanabe, et al in 2017 showed that the integrity of DNA stored in FFPE continued to decline. DNA integrity, which was measured by comparing the number of long and short segment amplicons, decreased significantly when compared at 6 months, 3 years and 12 years after the sample was taken. According to Berrino, et al., this problem can be ameliorated by modifying the tissue preservation technique using formalin at a cold temperature of 4oC. In this study, we also found that a greater proportion of CMV was detected in newer samples with a proportion reaching 30% in samples taken within 1 year before DNA extraction. In the bivariate analysis, we found that the CMV DNA positivity rate in this study was significantly higher in tissue samples obtained in the last three years compared to older tissue samples. The proportion of CMV DNA in our study, if only considering samples from the last three years, was 28.8%, higher than the results obtained by other studies [8, 15, 19, 20].

Data shows that Southeast Asia has a high CMV seroprevalence with positive rates of IgG seroprevalence against CMV ranging from 52.4% to 100% [2]. In Indonesia, there is no data regarding CMV seroprevalence in adults, but the CMV seroprevalence in newborns is 5.8%. In comparison, the global seroprevalence of CMV infection in newborns is 0.6 to 6%. This number put Indonesia in the higher range of CMV seroprevalence. This could be related to the relatively high proportion of CMV in our study, especially if only considering the

proportion of positive CMV DNA in the last three years' sample [21].

Research shows that CMV can be transmitted from various routes. The main route of entry for CMV into the body is through contact between infected body fluids such as saliva, breast milk, semen, tears, or urine with the mucosa. After entering the body, CMV can become latent or be carried through the bloodstream and infect various organs including the lungs. Cytomegalovirus can also be transmitted indirectly through contaminated surfaces. Some evidence shows that CMV DNA can be detected in the air of CMV-infected patient's rooms. This can explain how CMV can be detected in lung tissue [22, 23].

In this study, the tissue examined contained malignant cells from NSCLC patients. Compared to healthy tissue, the proportion of CMV DNA in this study was higher. In research published by Hazir-Konya, et al in 2020, CMV DNA was only found in 1.58% of gastrointestinal tissue samples. In healthy ovarian epithelial cells, CMV DNA was found at 0.5%. In the skin of healthy patients, CMV DNA was found in 5.1% of tissue samples. There is no data regarding the proportion of CMV in healthy lung tissue samples, but research by Lee shows a proportion of 18% in bronchial lavage samples. The higher proportion of CMV DNA in NSCLC tissue samples compared to healthy tissue could indicate that there is a possibility that CMV plays a role in the pathogenesis of NSCLC through mechanisms that are not yet clearly understood [24–27].

In this study, two genes were detected, the *MIE* and *gB* genes. The two primers for these genes were chosen based on previous research by Harabajsa, et al and have been proven to be able to detect traces of CMV DNA in lung cancer cells [8]. These two genes are highly conserved and can be found in all CMV strains because they have critical functions [28]. The *MIE* (major immediate early) gene is a gene involved in the virus replication cycle. The *MIE* gene is expressed very early even before the virus replication process takes place. Expression of the *MIE* gene will trigger the expression of other genes involved in the virus replication process. In latent CMV viruses, the *MIE* gene also plays a role in virus reactivation. The *gB* (glycoprotein B) gene is a gene encoding an envelope protein which plays a role in the process of virus entry into cells and the process of spreading CMV infection between cells. Research by Isaacson, et al shows that CMV which has a defect in *gB* gene expression can attach to cells but cannot enter the cell [29]. In cancer cells, the glycoprotein produced by the *gB* gene can interact with *EGFR* and activate the PI3K/AKT pathway so that increasing cell growth, cell survival and cancer cell migration [30, 31].

In this study, the *gB* gene was detected more frequently than the *MIE* gene. These results are different from the results of research by Harabajsa, et al which found more *MIE* genes but are similar to research conducted by Yang, et al on breast tumor samples. In this study, Yang, et al found that the *gB* gene was detected in 18.4% of breast tumor samples and no *MIE* gene was detected [32]. This difference could be caused by different CMV variants between population. A study shows that mutations in the *gB* gene cause CMV to enter cells through a different and faster process. It is reasonable to suspect that variants between different CMV strains could affect the DNA that can be detected by PCR. This difference may also be due to the fact that in these two studies, CMV could be in different phases of viral replication because the *MIE* and *gB* genes are expressed in two different phases of replication. Another possibility is that this difference could also be caused by differences in the types of samples studied. In the study by Harabajsa, the samples came from adenocarcinoma cytology samples, whereas in this study the samples came from paraffin block samples of NSCLC tissue which included adenocarcinoma and other types of NSCLC. Another possibility is that the differences could be caused by the small number of samples so that they do not represent real conditions. There are still few studies examining the *MIE* and *gB* genes in the same sample so further research is needed to find the implication of this difference [8, 32].

Bivariate analysis was done between CMV DNA and *EGFR* mutations. The *EGFR* mutations were detected more frequently in samples with the CMV gene, although this was not statistically significant. These results are in accordance with research by Harabajsa, et al which found that CMV DNA was significantly more common in lung cancer cytology samples with positive *EGFR* mutations [8]. It is known that CMV uses the *EGFR* receptor as a pathway for virus entry into cells [3]. Research by Cojohari, et al also showed that CMV can cause cells to enter an antiapoptotic state through binding of the *gB* glycoprotein on the viral envelope to the *EGFR* receptor on the cell surface through activation of the PI3K/AKT pathway [31].

The results of this study may provide interesting implications for either clinical practice or future research. These findings provide an indication that there may be a relationship between NSCLC and CMV and strengthen the suspicion that there is a role for CMV as an oncogenic virus. Further research needs to be carried out to determine the mechanisms underlying how CMV acts as an oncogenic agent. More sophisticated research methods need to be carried out, for example by DNA sequencing, to find out whether the CMV DNA has integrated with cancer cell DNA and caused mutations in the cancer cells or just a DNA contaminant in the tissue.

In clinical practice, these findings make it important for clinicians to remember that co-infection with CMV often occurs in cancer patients. In other studies conducted on post-surgical lung cancer and breast cancer, the detection of CMV DNA in breast cancer cells was associated with a worse prognosis [32, 33]. In research conducted by Stragliotto et al. [6] in 2013, it was shown that glioblastoma patients with CMV DNA detected in tissue samples have a significantly higher overall survival if they receive adjuvant antiviral therapy in the form of valganciclovir. If further research succeeds in proving that CMV is an oncogenic agent in NSCLC and antiviral administration is able to improve the prognosis of KPKBSK patients, then the management of KPKBSK can be improved for the better [6].

*Limitations*

This research has several limitations. In this study,

CMV DNA detection was done using FFPE samples taken within the last six years due to the limited number of FFPE samples that could be obtained. The quality of DNA in FFPE samples continues to decline over time that in older FFPE samples, a higher rate of false negatives can occur. In newer FFPE samples, a higher proportion of CMV DNA was detected. In future studies, samples obtained within the last three years should be used. Fresh samples should also be studied and compared with FFPE samples.

The methods used in this research were PCR and electrophoresis. This method is able to detect CMV DNA by using primers for the *MIE* gene and *gB* gene. However, the PCR and electrophoresis methods are unable to distinguish whether the CMV gene represent CMV coinfection in tissue samples, only a contaminant, or there has been integration between CMV DNA and cancer cell DNA. Another possibility that could provide similar results is that CMV does not play a role in oncogenesis but CMV DNA is detected because the immunosuppression process in cancer patients causes reactivation of latent CMV in the patient's body. Research shows that latent CMV infections are more susceptible to reactivation in cancer patients [34]. These questions can be confirmed if more sophisticated examination methods was used, for example by DNA sequencing (next gene sequencing) or using the fluorescence in situ hybridization (FISH) technique.

There are no universal standards regarding the minimum sample volume or an appropriate number of cells for CMV DNA examination. As a comparison, testing for *EGFR* mutations according to the literature requires at least 200 cells or 10% of the cells were tumor cells. In this study, FFPE samples that were considered adequate were samples that were assessed as adequate based on the observations by anatomical pathologists with a minimum cell count of 200 cells, referring to the minimum number of cells for molecular examination of *EGFR* mutations. However, another study by Harabjsa, et al which used cytology samples provided similar CMV DNA detection results. This result may reflect that the minimum sample volume required for CMV DNA detection is lower than that used in this study. In future studies, low volume of samples should not discourage researcher to test for CMV DNA. Further research is needed to determine the standard sample volume required for detection of CMV DNA in NSCLC tissue [35].

In conclusion, cytomegalovirus DNA can be detected in NSCLC FFPE samples and may indicate CMV's role in carcinogenesis. Further studies are needed to delineate its role in carcinogenesis.

## Author Contribution Statement

All authors contributed equally in this study.

## Acknowledgements

This research was partly funded by Respiratory and Tuberculosis Research and Training Center (SATURATE), Persahabatan Hospital, Jakarta, Indonesia.

*Approval*

This research was approved by Ethical Committee of Persahabatan Hospital, Jakarta Indonesia. Approval letter number 129/KEPK-RSUPP/09/2023. All subject's identities and data were kept confidential.

*Conflict of interest*

There is no conflict of interest in this research

## References

1. Poole EL, Nevels MM. Cytomegalovirus pathogenesis and host interactions. Frontiers in cellular and infection microbiology. 2021 Jul 7;11:711551.
2. Fuchs FE, Pauly M, Black AP, Hübschen JM. Seroprevalence of torch pathogens in southeast asia. Microorganisms. 2021;9(3):574. https://doi.org/10.3390/microorganisms9030574.
3. Falcão ASC, da Costa Vasconcelos PF, Lobato da Silva DF, Viana Pinheiro JJ, Falcão LFM, Quaresma JAS. Mechanisms of human cytomegalovirus infection with a focus on epidermal growth factor receptor interactions. Rev Med Virol. 2017;27(6):e1955. https://doi.org/10.1002/rmv.1955.
4. Krump NA, You J. Molecular mechanisms of viral oncogenesis in humans. Nat Rev Microbiol. 2018;16(11):684-98. https://doi.org/10.1038/s41579-018-0064-6.
5. Söderberg-Nauclér C. New mechanistic insights of the pathogenicity of high-risk cytomegalovirus (cmv) strains derived from breast cancer: Hope for new cancer therapy options. EBioMedicine. 2022;81:104103. https://doi.org/10.1016/j.ebiom.2022.104103.
6. Stragliotto G, Rahbar A, Solberg NW, Lilja A, Taher C, Orrego A, et al. Effects of valganciclovir as an add-on therapy in patients with cytomegalovirus-positive glioblastoma: A randomized, double-blind, hypothesis-generating study. Int J Cancer. 2013;133(5):1204-13. https://doi.org/10.1002/ijc.28111.
7. Ahmed HG, Suliman RSA, Ashankyty IM, Albieh ZA, Warille AA. Role of human cytomegalovirus in the etiology of nasopharyngeal carcinoma. J Cancer Res Ther. 2018;14(3):583-6. https://doi.org/10.4103/0973-1482.176175
8. Harabajsa S, Šefčić H, Klasić M, Milavić M, Židovec Lepej S, Grgić I, et al. Infection with human cytomegalovirus, epstein-barr virus, and high-risk types 16 and 18 of human papillomavirus in egfr-mutated lung adenocarcinoma. Croat Med J. 2023;64(2):84-92. https://doi.org/10.3325/cmj.2023.64.84.
9. Laporan nasional RISKESDAS 2018. Jakarta: Kementerian Kesehatan RI; 2019.
10. Garrison GW, Cho JL, Deng JC, Camac E, Oh S, Sundar K, et al. Ats core curriculum 2021. Adult pulmonary medicine: Thoracic oncology. ATS Sch. 2021;2(3):468-83. https://doi.org/10.34197/ats-scholar.2021-0032RE.
11. Siddiqui F, Vaqar S, Siddiqui AH. Lung Cancer. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 [cited 2023 Mar 18].
Available from: http://www.ncbi.nlm.nih.gov/books/NBK482357/
12. Andarini S, Syahruddin E, Aditya N, Zaini J, Kurniawan FD, Ermayanti S, et al. Indonesian Society of Respirology (ISR) Consensus Statement on Lung Cancer Screening and Early Detection in Indonesia. J Respirologi Indones. 2023;43(2):144–50. https://doi.org/10.36497/jri.v43i2.455
13. Dickerson F, Katsafanas E, Origoni A, Newman T, Rowe K, Ziemann RS, et al. Cigarette smoking is associated with

Herpesviruses in persons with and without serious mental illness. PLOS ONE. 2023;18(1):e0280443. https://doi.org/10.1371/journal.pone.0280443
14. Jusuf A, Wibawanto A, Icksan AG, Syahruddin E, Juniarti ES. Lung Cancer: Guidelines for diagnosis and management in Indonesia. Jakarta: Universitas Indonesia Publisher; 2018.
15. Zhang H, Ehiri J, Yang H, Tang S, Li Y. Impact of community-based dot on tuberculosis treatment outcomes: A systematic review and meta-analysis. PLoS One. 2016;11(2):e0147744. https://doi.org/10.1371/journal.pone.0147744.
16. Han B, Tjulandin S, Hagiwara K, Normanno N, Wulandari L, Laktionov K, et al. Egfr mutation prevalence in asia-pacific and russian patients with advanced nsclc of adenocarcinoma and non-adenocarcinoma histology: The ignite study. Lung Cancer. 2017;113:37-44. https://doi.org/10.1016/j.lungcan.2017.08.021.
17. Syahruddin E, Wulandari L, Sri Muktiati N, Rima A, Soeroso N, Ermayanti S, et al. Uncommon egfr mutations in cytological specimens of 1,874 newly diagnosed indonesian lung cancer patients. Lung Cancer (Auckl). 2018;9:25-34. https://doi.org/10.2147/lctt.S154116.
18. Li AR, Chitale D, Riely GJ, Pao W, Miller VA, Zakowski MF, et al. Egfr mutations in lung adenocarcinomas: Clinical testing experience and relationship to egfr gene copy number and immunohistochemical expression. J Mol Diagn. 2008;10(3):242-8. https://doi.org/10.2353/jmoldx.2008.070178.
19. Berrino E, Annaratone L, Miglio U, Maldi E, Piccinelli C, Peano E, et al. Cold formalin fixation guarantees DNA integrity in formalin fixed paraffin embedded tissues: Premises for a better quality of diagnostic and experimental pathology with a specific impact on breast cancer. Front Oncol. 2020;10:173. https://doi.org/10.3389/fonc.2020.00173.
20. Watanabe M, Hashida S, Yamamoto H, Matsubara T, Ohtsuka T, Suzawa K, et al. Estimation of age-related DNA degradation from formalin-fixed and paraffin-embedded tissue according to the extraction methods. Exp Ther Med. 2017;14(3):2683-8. https://doi.org/10.3892/etm.2017.4797.
21. Putri ND, Wiyatno A, Dhenni R, Sriyani IY, Dewantari AK, Handryastuti S, et al. Birth prevalence and characteristics of congenital cytomegalovirus infection in an urban birth cohort, jakarta, indonesia. Int J Infect Dis. 2019;86:31-9. https://doi.org/10.1016/j.ijid.2019.06.009.
22. McCluskey R, Sandin R, Greene J. Detection of airborne cytomegalovirus in hospital rooms of immunocompromised patients. J Virol Methods. 1996;56(1):115-8. https://doi.org/10.1016/0166-0934(95)01955-3.
23. Sayahi T, Nielson C, Yu Y, Neuberger K, Seipp M, Firpo MA, et al. Airborne aerosolized mouse cytomegalovirus from common otolaryngology procedures: Implications for covid-19 infection. Otolaryngol Head Neck Surg. 2021;164(3):547-55. https://doi.org/10.1177/0194599820957966.
24. Hazır-Konya H, Avkan-Oğuz V, Akpınar H, Sağol Ö, Sayıner A. Investigation of cytomegalovirus in intestinal tissue in a country with high cmv seroprevalence. Turk J Gastroenterol. 2021;32(2):123-32. https://doi.org/10.5152/tjg.2021.191008..
25. Ingerslev K, Høgdall E, Skovrider-Ruminski W, Schnack TH, Lidang M, Høgdall C, et al. The prevalence of ebv and cmv DNA in epithelial ovarian cancer. Infect Agent Cancer. 2019;14:7. https://doi.org/10.1186/s13027-019-0223-z.
26. Ohtsuka T, Yamazaki S. Prevalence of human cytomegalovirus DNA in scleroderma skin tissue. Int J Dermatol. 2006;45(1):86-7. https://doi.org/10.1111/j.1365-4632.2004.02576.x.
27. Lee HY, Rhee CK, Choi JY, Lee HY, Lee JW, Lee DG. Diagnosis of cytomegalovirus pneumonia by quantitative polymerase chain reaction using bronchial washing fluid from patients with hematologic malignancies. Oncotarget. 2017;8(24):39736-45. https://doi.org/10.18632/oncotarget.14504.
28. Zweygberg Wirgart B, Brytting M, Linde A, Wahren B, Grillner L. Sequence variation within three important cytomegalovirus gene regions in isolates from four different patient populations. J Clin Microbiol. 1998;36(12):3662-9. https://doi.org/10.1128/jcm.36.12.3662-3669.1998.
29. Isaacson MK, Compton T. Human cytomegalovirus glycoprotein b is required for virus entry and cell-to-cell spread but not for virion attachment, assembly, or egress. J Virol. 2009;83(8):3891-903. https://doi.org/10.1128/jvi.01251-08.
30. Cobbs C, Khan S, Matlaf L, McAllister S, Zider A, Yount G, et al. Hcmv glycoprotein b is expressed in primary glioblastomas and enhances growth and invasiveness via pdgfr-alpha activation. Oncotarget. 2014;5(4):1091-100. https://doi.org/10.18632/oncotarget.1787.
31. Cojohari O, Peppenelli MA, Chan GC. Human cytomegalovirus induces an atypical activation of akt to stimulate the survival of short-lived monocytes. J Virol. 2016;90(14):6443-52. https://doi.org/10.1128/jvi.00214-16.
32. Yang Z, Tang X, Hasing ME, Pang X, Ghosh S, McMullen TPW, et al. Human cytomegalovirus seropositivity and viral DNA in breast tumors are associated with poor patient prognosis. Cancers (Basel). 2022;14(5). https://doi.org/10.3390/cancers14051148
33. Cortón Ruiz J, Lücke L, Minarowski Ł, Kowalczuk O. Human cytomegalovirus infection effects on lung cancer prognosis after surgical resection. Eur Respir J. 2019;54(suppl 63):PA3053. https://doi.org/10.1183/13993003.congress-2019.PA3053
34. Schlick K, Grundbichler M, Auberger J, Kern JM, Hell M, Hohla F, et al. Cytomegalovirus reactivation and its clinical impact in patients with solid tumors. Infect Agent Cancer. 2015;10:45. https://doi.org/10.1186/s13027-015-0039-4
35. Haralsingh A, West M. Tissue yields for epidermal growth factor receptor analysis in non-small cell lung cancer patients in trinidad and tobago. Cureus. 2021;13(1):e12531. https://doi.org/10.7759/cureus.12531.