

1     **Association Between Somatic *PIK3CA* Mutations and Ancestry in**  
2             **South Asia: Prognostic Insights From a Sri Lankan Breast**  
3                     **Cancer Study**

4     Tharini Ruwinya Cabraal<sup>1\*</sup>, Iranthi Kumarasinghe<sup>2¶</sup>, Ranga Perera<sup>2¶</sup>, Jayantha Balawardane<sup>2¶</sup>,  
5             Sameera Viswakula<sup>3¶</sup>, Nandika Perera<sup>4¶</sup>, Gayani Harendra Galhena<sup>1\*</sup>

6             1. Centre for Immunology and Molecular Biology, Department of Zoology &  
7             Environmental Sciences, Faculty of Science, University of Colombo, Colombo, Sri  
8                     Lanka

9             2. University Hospital, General Sir John Kotelawala Defence University, Colombo, Sri  
10                     Lanka

11             3. Department of Statistics, Faculty of Science, University of Colombo, Colombo, Sri  
12                     Lanka

13             4. Genetech Molecular Diagnostics, Colombo, Sri Lanka

15     **\*Corresponding Authors:**

16     Prof. Gayani Harendra Galhena

17     Email: [gayani@zoology.cmb.ac.lk](mailto:gayani@zoology.cmb.ac.lk) (GHG)

18     ORCID: <https://orcid.org/0000-0002-1617-6612>

19

20     Tharini Ruwinya Cabraal

21     Email: [tharinicabraal@gmail.com](mailto:tharinicabraal@gmail.com)

22     ORCID: <https://orcid.org/0009-0003-8943-0110>

23

24     ¶ These authors contributed equally to this work.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## 25 **Abstract (299 words)**

### 26 **Background**

27 The *PIK3CA* oncogene is one of the most mutated oncogenes in breast cancer, with ancestry-  
28 linked variations reported globally. The study aimed to discern the prevalence and prognostic  
29 role of *PIK3CA* mutations in Sri Lankan breast cancer patients for the first time, focusing on  
30 the correlation between somatic mutations and patient ancestry in an exclusively South Asian  
31 cohort of breast cancers.

### 32 **Materials & Methods**

33 A qPCR-based genetic analysis was performed on DNA from formalin-fixed, paraffin-  
34 embedded (FFPE) tissue samples of 63 clinically diagnosed, female, Sri Lankan breast cancer  
35 patients using the QClamp® *PIK3CA* Mutation Detection Test for the hotspot mutations of  
36 *PIK3CA* (i.e., H1047R, E545K, E542K) followed by a statistical analysis. Patient samples  
37 and clinical data were fully anonymized, with no identifying information available to authors  
38 at any point of the study.

### 39 **Results**

40 Somatic missense *PIK3CA* mutations H1047R and E542K, were detected in 17.46% of the  
41 cohort. E545K mutation was not detected. The observed mutations were associated with an  
42 increased risk of lymph node (LN) metastasis ( $p=0.036$ , OR 9.60) and reduced recurrence-  
43 free survival (RFS) ( $p<0.001$ , HR 26.19). In addition, LN metastasis ( $p=0.026$ , HR 123.94)  
44 and a high Ki67 index ( $p=0.029$ , HR 79.69) were individually associated with reduced RFS.  
45 All three factors above- presence of a *PIK3CA* mutation, LN metastasis and a high Ki67  
46 index- in combination, were also associated with reduced RFS ( $p<0.001$ ). Further analysis  
47 revealed a significant association between patient ancestry and *PIK3CA* mutation status  
48 ( $p=0.028$ ).

### 49 **Conclusion**

50 Despite being a pilot study, the findings suggest that *PIK3CA* may serve as a potential  
51 prognostic biomarker in Sri Lankan breast cancer patients, with ancestry-linked variations  
52 potentially influencing metastatic outcomes. These results bring out the importance of  
53 integrating PI3K inhibitors into the therapeutic management of breast cancers in Sri Lanka  
54 after validating these findings in a functional study using a larger cohort.

## 55 Introduction

56 Breast cancer is the most commonly diagnosed cancer in women worldwide [1]. Asian women,  
57 particularly those in developing countries, often have a poorer prognosis for breast cancer, with  
58 higher cancer grades and stages as well as an earlier onset of the disease [2].

59 Breast cancer has distinct molecular subtypes that are categorized based on the  
60 immunohistochemical (IHC) expression of steroid hormone receptors (SHR), including  
61 estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor  
62 receptor 2 (HER2/*ERBB2* or HER2/*neu*). These subtypes are Luminal A, Luminal B, HER2-  
63 positive (HER2+), and triple-negative breast cancer (TNBC) [1].

64 More than 85% of breast cancers are sporadic, primarily caused by somatic mutations [3].  
65 These mutations in oncogenes (OGs) and tumor suppressor genes (TSGs) often initiate  
66 tumorigenesis [4]. Breast cancer tumorigenesis has been linked to mutations in several cell  
67 signaling pathways crucial for normal development such as the PI3K/AKT/mTOR pathway,  
68 commonly referred to as the PI3K pathway. Genetic hyperactivation of this pathway is one of  
69 the most common driver mechanisms in breast cancer, often through alterations to its  
70 components like p110 $\alpha$  (encoded by the phosphatidylinositol-4,5-bisphosphate 3-kinase  
71 catalytic subunit alpha gene/*PIK3CA*; *NCBI Gene ID 5290*), p110 $\beta$ , and p85 $\alpha$  [1,5,6].

72 Approximately 30% of the mechanisms that over-activate PI3K signaling are due to somatic  
73 mutations in the *PIK3CA* oncogene [5], making it the most commonly mutated oncogene in  
74 breast cancer [6-8]. While *PIK3CA* mutations can occur across the p110 $\alpha$  subunit, about 80%  
75 are caused by the three hotspot mutations clustered in two domains of p110 $\alpha$ : E545K and  
76 E542K in the helical domain encoded by exon 9, and H1047R in the kinase domain encoded

77 by exon 20. Among these, H1047R is the most frequently detected *PIK3CA* mutation, while  
78 E542K is the least prevalent [9].

79 Studies have shown that between 10 to 45% of all human breast cancers harbor *PIK3CA* driver  
80 mutations, with the exact percentage varying based on the studied population and ancestry [5-  
81 10]. Activating mutations of *PIK3CA* are reported in nearly 42% of SHR+/HER2-, 31% of  
82 HER2+ and 16% of TNBC tumors [5].

83 Importantly, the presence of *PIK3CA* mutations has been associated with poorer prognostic  
84 outcomes, especially in ER+ cases [11]. Numerous studies illustrate that *PIK3CA* mutations  
85 are linked with resistance to various therapies, including endocrine therapy, chemotherapy,  
86 radiotherapy, anti-HER2 therapy and immunotherapy. This resistance negatively affects all  
87 available treatment modalities for all breast cancer subtypes, with more pronounced effects in  
88 advanced and metastatic cases [5,7-10,12]. These mutations also promote cell division and  
89 inhibit apoptosis in TNBC cells, leading to resistance to their primary treatment option:  
90 chemotherapy [12]. In addition to their role in treatment resistance, accumulating evidence  
91 suggests a strong association between *PIK3CA* mutations and tumor recurrence, poor  
92 relapse/recurrence-free survival (RFS), and/or overall survival (OS), particularly in  
93 postoperative Asian patients with metastatic SHR+/HER2- breast cancers [9,11,13-15].  
94 Despite extensive research, the overall relationship between *PIK3CA* mutations and the  
95 clinicopathological profile of breast cancer patients remains controversial, with conflicting  
96 findings across various patient populations, such as observed associations of *PIK3CA*  
97 mutations with longer RFS and/or OS [16,17].

98 Recent advances in genomic studies have provided significant insights into the relationship  
99 between genetic ancestry and somatic mutations across cancer types via changes in the  
100 germline [18-22]. In particular, the *PIK3CA* oncogene is reported as one of the top two genes

101 observed in breast cancers, whose somatic mutations have shown significant ancestry-linked  
102 differences in several global studies predominantly involving patients from European and  
103 African ancestries [19-21]. However, Asian populations, particularly those of South Asian  
104 descent, remain highly underrepresented in these studies, limiting a comprehensive  
105 understanding of ancestry-driven somatic mutations in the region [22].

106 Sri Lanka is a unique island nation in the Indian Ocean with a rich tapestry of ethnic diversity.  
107 Over 70% of the population is composed of Sinhalese, descendants of Indo-Aryan ancestry,  
108 found exclusively in Sri Lanka [23]. Other major ethnic groups include Tamils and Moors, who  
109 trace their roots to Dravidian ancestry, and Burghers, descendants of European colonists in Sri  
110 Lanka, who also share Indo-Aryan ancestry [23,24]. This distinct combination of Indo-Aryan  
111 and Dravidian ancestries, together with its centralized geographical location, particularly  
112 during the maritime silk route, makes Sri Lanka a melting pot of genetic diversity.

113 Breast cancer is the most frequently diagnosed cancer among Sri Lankan women [25]. As a  
114 developing, lower-middle income island nation in the Indian Ocean, Sri Lanka has limited  
115 research on breast cancer trends and genetic markers, despite a significant rise in incidence and  
116 mortality rates over time [26,25]. To the best of our knowledge, no studies have been conducted  
117 on the prevalence and prognostic outcomes of *PIK3CA* mutations and their correlations with  
118 ancestry in breast cancer patients of Sri Lanka. Given the mutational spectrum of *PIK3CA* and  
119 its impact on the clinicopathological profile and prognosis of breast cancer, which varies  
120 significantly depending on the population and their ancestry, it is important to elucidate the  
121 current status of Sri Lankan breast cancer patients in this context – an area that remains largely  
122 unexplored.

123 This study was therefore designed to achieve the overarching aim of cataloguing hotspot  
124 mutations in the *PIK3CA* oncogene within a cross-section of Sri Lankan breast cancer patients

125 that includes all different ancestral groups. Using DNA extracted from tumor tissues obtained  
126 at the point of diagnosis, we aimed to correlate the detected mutations with specific  
127 clinicopathological characteristics and ultimately assess the prognostic potential of *PIK3CA*  
128 mutations using RFS as a key metric. Given that the Sinhalese population is found nowhere  
129 else in the world (except for few migrant individuals), results from this study offer a rare  
130 opportunity to explore how ancestral genetic makeup in an island nation intersects with somatic  
131 mutations, such as those in the *PIK3CA* oncogene. This study managed to successfully catalog  
132 *PIK3CA* mutations in our cohort of Sri Lankan breast cancer patients and deduced their impacts  
133 on disease prognosis while addressing the critical gap on the implications of ancestry on  
134 somatic mutations in a purely South Asian cohort of breast cancer patients, for the first time.

## 135 **Materials and methods**

### 136 **Patient samples**

137 Ethical approval for the analysis of an existing set of formalin-fixed paraffin-embedded (FFPE)  
138 breast tumor biopsy samples and fully anonymized medical records of the relevant patients was  
139 obtained from the Ethics Review Committee of the Institute of Biology, Sri Lanka (ERC  
140 IOBSL 303 07 2023, S1 Fig). This retrospective study was conducted in compliance with the  
141 REMARK (Reporting Recommendations for Tumor Marker Prognostic Studies) and ICMJE  
142 (International Committee of Medical Journal Editors) guidelines.

143 The fully anonymized database of breast cancer patient records from the Diagnostics  
144 Laboratory at the University Hospital of General Sir John Kotelawala Defense University was  
145 accessed on the 27<sup>th</sup> of August 2023. All patient data were fully anonymized, ensuring no  
146 patient identifiers were available to the authors at any stage of the research. Using this database,  
147 medical records of breast cancer patients treated at the University Hospital of General Sir John

148 Kotelawala Defense University from 2019 to 2023, who met the inclusion criteria (Fig 1) were  
149 selected for the study. The corresponding FFPE blocks were retrieved from the hospital's  
150 diagnostics laboratory and stored at room temperature until further analysis.  
151 Clinicopathological data and follow-up information of these patients ranging from 1 to 5 years  
152 post-diagnosis were collected from the hospital database.

153 To ensure patient anonymity, medical and follow-up data of individual patients were recorded,  
154 and the samples were identified exclusively using identification numbers assigned to each  
155 sample for the purpose of research. Therefore, as authors did not have access to information  
156 that could identify individual participants during or after data collection, total anonymity of  
157 patient identity was maintained throughout the study.

158 **Fig 1. Flowchart of patient enrollment, inclusion criteria, and study workflow**

## 159 **DNA extraction and hotspot mutation assay**

160 Genomic DNA was extracted from FFPE samples using the QIAamp DNA FFPE Tissue Kit  
161 according to the manufacturer's instructions (Cat. No. 56404; Qiagen, Hilden, Germany) and  
162 assessed for quality and quantity using the Thermo Scientific™ NanoDrop™ One  
163 Spectrophotometer. The three hotspot mutations of the *PIK3CA* oncogene- H1047R, E545K,  
164 and E542K- were genotyped in 63 of the selected samples using the QClamp® *PIK3CA*  
165 Mutation Detection Test Kit (Cat. No. DC-10-1072R; DiaCarta, CA 94588, USA), which  
166 utilized a Xeno-nucleic acid (XNA)-based mutation-specific quantitative real-time polymerase  
167 chain reaction (qPCR) method.

## 168 **Endpoint and statistical analysis**

169 The prevalence of *PIK3CA* mutations in the study cohort and their associations with  
170 clinicopathological features were determined using the  $\chi^2$  test. Fisher's exact test was

171 employed for dichotomous features. To enhance the statistical validity and minimize the risk  
172 of type I errors, patients were grouped according to significant, widely accepted  
173 clinicopathological cutoffs in oncology based on the sample size available for each category  
174 (S1 Table). In the analysis of ancestry, patients grouped into Indo-Aryan ancestry (Sinhalese  
175 & Burghers) were compared to those of Dravidian ancestry (Tamils & Moors). Depending on  
176 data normality, either a t-test or Mann-Whitney Test was performed to compare the mean or  
177 median of populations for the relevant variables. For specific variables showing significant  
178 associations, binary logistic regression was used to obtain the odds ratio (OR) and the 95%  
179 confidence interval (CI). This binary logistic regression model assumed that there were no  
180 significant multicollinearity issues, and all independent variables were linearly related to the  
181 log-odds of the dependent variable.

182 All patients underwent routine clinical examinations at the hospital every 3-6 months, and the  
183 clinicopathological variables analyzed were those recorded at the initial diagnosis. Since all  
184 patients with survival data (n=42) were alive at the last follow-up, RFS was used as the survival  
185 metric instead of OS, and was defined as the time from the date of curative surgery to the  
186 clinical diagnosis of recurrence. Events considered in RFS included distant and local invasive  
187 recurrence. Patients without survival data were excluded from the survival analysis. The  
188 median follow-up time after diagnosis was 37.00 months (0.00–81.19).

189 The Kaplan-Meier method was used to plot RFS, while the Mantel-Cox log-rank test was  
190 applied to compare different subgroups of the variables. Significant parameters identified by  
191 the log-rank test ( $p < 0.05$ ) were further analyzed using a univariate Cox proportional hazard  
192 regression model. A multivariate model was finally used to analyze the combined effect of all  
193 significant associations on RFS. However, for parameters which could not generate a



194 statistically robust Cox proportional hazard regression model due to limited data on survival,  
195 only the p-value was reported.

196 Samples with missing data for the variable of interest were excluded from specific analyses.  
197 All statistical tests were two-sided and a p-value below 0.05 was considered statistically  
198 significant. Analyses were performed using SPSS 25 (Released 2017. IBM SPSS Statistics for  
199 Windows, Version 25.0. Armonk, NY: IBM Corp.) and R (version 4.0.2, R Foundation for  
200 Statistical Computing, Vienna, Austria).

## 201 **Results**

### 202 **Analysis of the clinicopathological features and ethnicity**

203 The studied cohort consisted of 66 patients clinically diagnosed with breast cancer between  
204 2019 and 2023. The median age at the time of diagnosis was 58.00 years (S2 Fig).

205 Table 1 presents several key characteristics of the cohort. Accordingly, IHC analysis revealed  
206 that over 80% of the tumors were SHR-positive HER2-negative luminal subtypes, with luminal  
207 A and B subtypes occurring in equal proportions. Over 95% of the cohort comprised of  
208 Sinhalese patients while the remaining 5% included two Tamil patients, one Moor, and one  
209 Burgher patient. This represents only a rough approximation of the actual ethnic distribution  
210 in Sri Lanka, where approximately 72% of the population are Sinhalese [23].

211 **Table 1. Clinicopathological and population characteristics of the cohort**

212

<b>Age at diagnosis</b>	<b>No. of patients</b>	<b>%</b>
<35	0	0.00
36-45	5	7.58
46-55	20	30.30
56-65	19	28.79
66-75	15	22.73
76-85	6	9.09
>85	1	1.52
<b>Side</b>	<b>No. of patients</b>	<b>%</b>

Table 1 continued

L/S	30	45.45
R/S	36	54.55
<b>Site of the tumor</b>	<b>No. of patients</b>	<b>%</b>
Upper outer quadrant	2	3.03
Upper inner quadrant	14	21.21
Lower inner quadrant	0	0.0
Lower outer quadrant	3	4.55
Central	5	7.58
Clock Positions	22	33.33
Two quadrants	2	3.03
Unknown	18	27.27
<b>Histological type</b>	<b>No. of patients</b>	<b>%</b>
IBC-NST	57	86.36
Invasive mucinous carcinoma	4	6.06
<b>Histological type</b>	<b>No. of patients</b>	<b>%</b>
Invasive lobular carcinoma	2	3.03
Invasive papillary carcinoma	1	1.52
Invasive tubular carcinoma	1	1.52
DCIS	1	1.52
<b>Tumor grade/ differentiation</b>	<b>No. of patients</b>	<b>%</b>
Grade 1/ well-differentiated	22	33.34
Grade 2/ moderately differentiated	37	56.06
Grade 3/ poorly differentiated	5	7.57
Unknown	2	3.03
<b>Tumor size</b>	<b>No. of patients</b>	<b>%</b>
≤2 cm	12	18.18
>2 cm	37	56.06
Unknown	17	25.76
<b>Pathological tumor stage</b>	<b>No. of patients</b>	<b>%</b>
Tis	1	1.52
pT1	10	15.15
pT2	31	46.97
pT3	6	9.09
Unknown	18	27.27
<b>Pathological nodal stage</b>	<b>No. of patients</b>	<b>%</b>
pN0	29	43.94
pN1	13	19.70
<b>Pathological nodal stage</b>	<b>No. of patients</b>	<b>%</b>
pN2	6	9.09
pN3	2	3.03
Unknown	16	24.24
<b>Lymph node metastasis</b>	<b>No. of patients</b>	<b>%</b>
Positive	21	31.82
Negative	29	43.94
Unknown	16	24.24
<b>ER status</b>	<b>No. of patients</b>	<b>%</b>

213

214

Table 1 continued

ER+	60	90.91
ER-	6	9.09
<b>PR status</b>	No. of patients	%
PR+	46	69.70
PR-	20	30.30
<b>SHR status</b>	No. of patients	%
SHR+	60	90.91
SHR-	6	9.09
<b>HER2 status</b>	No. of patients	%
HER2+	6	9.09
HER2-	59	89.39
Unknown	1	1.52
<b>Receptor based subtype</b>	No. of patients	%
SHR+/HER2-	55	83.34
SHR+/HER2+	4	6.06
<b>Receptor based subtype</b>	No. of patients	%
SHR-/HER2+	2	3.03
SHR-/HER2-	4	6.06
Unknown	1	1.51
<b>Ki67 Proliferation index</b>	No. of patients	%
<20% (low)	30	45.45
≥20% (high)	28	42.42
Unknown	8	12.12
<b>IHC-based molecular subtype</b>	No. of patients	%
Luminal A	24	36.36
Luminal B	24	36.36
Luminal (undetermined for A/B as Ki67 data unavailable)	7	10.60
HER2 enriched	6	9.09
TNBC	4	6.06
HER2 unknown	1	1.51
<b>Ethnicity</b>	No. of patients	%
Sinhalese	63	95.45
Tamil	1	1.51
Moor	1	1.51
Burgher	1	1.51
<b>Ancestry</b>	No. of patients	%
Indo-Aryan	64	96.96
Dravidian	2	3.03

DCIS: ductal carcinoma in-situ, ER: estrogen receptor, HER2: human epidermal growth factor receptor 2, IBC-NST: invasive breast carcinoma of no special type, IHC:immunohistochemical, L/S: left side, pN: pathological nodal stage, PR: progesterone receptor, pT: pathological tumor stage, R/S: Right side, SHR: steroid hormone receptor, Tis: carcinoma in-situ stage, TNBC: triple-negative breast cancer  
“Unknown” - samples with data unavailable for the parameter of interest

215

216

## 217 **Prevalence of *PIK3CA* mutations in the cohort**

218 Missense mutations in the *PIK3CA* oncogene were detected in 17.46% (n=11) of the cohort.  
219 The mutations observed included H1047R (c.3140A>G) in 81.81% (n=9) and E542K  
220 (c.1624G>A) in 18.18% (n=2) of the mutated samples. Notably, both H1047R and E542K  
221 mutations were detected in one sample (co-mutation prevalence: 9.09% of the mutated and  
222 1.58% of the total samples). No E545K mutations (c.1633G>A) were detected in the cohort.

## 223 **Association of *PIK3CA* mutations with clinicopathological** 224 **parameters and ancestry**

225 All *PIK3CA* mutations identified were found in SHR+/HER2- luminal tumors (Table 2).  
226 Importantly, a significant association was observed between *PIK3CA* mutational status and  
227 lymph node (LN) metastasis, with patients harboring a *PIK3CA* mutation exhibiting 9.60 times  
228 higher odds of LN metastasis compared to those without a *PIK3CA* mutation (p=0.036, OR  
229 9.60, 95% CI 1.05-87.78). Notably, a significant association was observed between the  
230 ancestry of the patients and the presence of a *PIK3CA* mutation (p=0.028). Other parameters  
231 showed no significant association with the mutational status.

232 **Table 2. *PIK3CA* mutational status and Ki67 proliferative index in selected**  
233 **clinicopathological and ancestral groups**

234

Parameter	Mutations absent (n)	<i>PIK3CA</i> Mutated (n)	Total (n)	p-value
All	52	11	63	
<b>Age at diagnosis</b>				
≤ 50 years	14	2	16	
> 50 years	38	9	47	0.714
<b>Histological type</b>				
IBC-NST	45	10	55	
Other	7	1	8	1.000

Table 2 continued

<b>Tumor grade/ differentiation</b>				
G1/ well-differentiated	17	4	21	
G2 or G3/ moderately or poorly differentiated	33	7	40	1.000
<b>Tumor size</b>				
≤ 2 cm	9	2	11	
> 2 cm	30	5	35	1.000
Mean (cm) (95% CI)	3.5 (2.8-4.2)	3.2 (1.7-4.7)		
Median (cm) (Range)	3.2 (0.7-12.0)	3 (1.1-6.0)		0.994
<b>Pathological tumor stage</b>				
Early stage (pT1/pT2)	32	6	38	
Late stage (pT3/pT4)	6	1	7	1.000
<b>Lymph node Metastasis</b>				
Negative	24	1	25	
Positive	15	6	21	<b>*0.036</b>
<b>ER status</b>				
ER+	47	11	58	
ER-	5	0	5	0.576
<b>PR status</b>				
PR+	37	7	44	
PR-	15	4	19	0.721
<b>SHR status</b>				
SHR+	47	11	58	
SHR-	5	0	5	0.576
<b>HER2 status</b>				
HER2+	6		6	
HER2-	45	11	56	0.580
<b>Receptor based subtype</b>				
SHR+/HER2-	42	11	53	
Other	10	0	10	0.187
<b>Ki67 Proliferation index</b>				
<20%	23	5	28	
≥20%	22	5	27	1.000
<b>IHC-based molecular subtype</b>				
Luminal	42	11	53	
HER2/TNBC	9	0	9	0.341
<b>Ancestry</b>				
Indo-Aryan	52	9	61	

235

Table 2 continued

Dravidian	0	2	2	<b>*0.028</b>
<b>Parameter</b>	<b>Ki67 &lt; 20% (Low)</b>	<b>Ki67 ≥ 20% (High)</b>	<b>Total</b>	<b>p-value</b>
All	30	28	58	
<b>Age at diagnosis</b>				
≤50 years	7	7	14	
>50 years	23	21	44	1.000
<b>Histological type</b>				
IBC-NST	23	27	50	
Other	7	1	8	0.053
<b>Tumor grade/ differentiation</b>				
G1/ well- differentiated	14	5		
G2 or G3/ moderately or poorly differentiated	15	23	38	<b>▲ 0.024</b>
<b>Tumor size</b>				
≤ 2 cm	5	5	10	
> 2 cm	20	12	32	0.714
Mean (cm) (95% CI)	3.3 (2.5-4.2)	3.4 (2.3- 4.5)		
Median (cm) (Range)	3.0 (1.2-12.0)	3.3 (0.7- 9.3)		0.884
<b>Pathological tumor stage</b>				
Early stage (pT1/pT2)	21	14	35	
Late stage (pT3/pT4)	2	3	5	0.634
<b>Lymph node metastasis</b>				
Negative	17	9	26	
Positive	8	8	16	0.353
<b>ER status</b>				
ER+	27	26	53	
ER-	3	2	5	1.000
<b>PR status</b>				
PR+	23	18	41	
PR-	7	10	17	0.390
<b>SHR status</b>				
SHR+	27	26	53	
SHR-	3	2	5	1.000
<b>HER2 status</b>				
HER2+	3	3	6	

Table 2 continued

HER2-	26	25	51	1.000
<b>Receptor based subtype</b>				
SHR+/HER2-	24	24	48	
Other	6	4	10	0.732
<b>IHC-based molecular subtype</b>				
Luminal	24	24	48	
HER2/TNBC	5	4	9	1.000

\*: association is significant at 5% level, OR 9.600, 95% CI 1.050-87.784

▲ : association is significant at 5% level, OR 4.293, 95% CI 1.279-14.408

CI: confidence interval, ER: estrogen receptor, HER2: human epidermal growth factor receptor 2, IBC-NST: invasive breast carcinoma of no special type, IHC: immunohistochemical, LN: Lymph Node, OR: odds ratio, PR: progesterone receptor, pT: pathological tumor stage, SHR: steroid hormone receptor, TNBC: triple-negative breast cancer

236

## 237 **Associations among other clinical and histopathological** 238 **parameters**

239 Table 2 also outlines the associations between the Ki67 proliferation index and other  
240 clinicopathological features. A significant association was identified between Ki67 status and  
241 tumor grade, with higher Ki67 proliferative indices being linked to a 4.29 times higher  
242 likelihood of developing moderately or poorly differentiated breast cancers of higher tumor  
243 grades, compared to those with proliferative indices < 20% (p=0.024, OR 4.29, 95% CI 1.28-  
244 14.41).

245 No statistically significant associations were observed between LN metastasis or patient age  
246 group and other clinicopathological parameters (S2 Table).

## 247 **Impact of *PIK3CA* mutational status on RFS**

248 Among patients with survival data, 9.5% (n=4) experienced a relapse in a different organ (e.g.,  
249 pelvic bone and lungs) following curative surgery. Three of these patients (75.0%) had *PIK3CA*  
250 mutations, with two harboring the H1047R mutation (50%) and one carrying the E542K  
251 mutation (25%).

252 The presence of a *PIK3CA* mutation was thereby significantly associated with reduced RFS  
 253 (p<0.001), as shown in the Kaplan-Meier curve (Fig 2A and Table 3). Patients with a *PIK3CA*  
 254 mutation had 26.19 times higher odds of relapse compared to the those without a mutation (HR  
 255 26.19, 95% CI 2.67-256.47). The mean RFS for patients with *PIK3CA* mutations was 26.5  
 256 months.

257 **Fig 2. Kaplan-Meier survival estimates for relapse-free survival (RFS) stratified by**  
 258 ***PIK3CA* mutation status (A), status of lymph node (LN) metastasis (B) and Ki67**  
 259 **proliferative status (C) while (D) shows the multivariate impact of all above three**  
 260 **factors – *PIK3CA* mutation, LN metastasis and Ki67 index on RFS.**

261 Fig. 2A: p-value<0.001, HR 26.187, 95% CI 2.674 – 256.475

262 Fig. 2B: p-value=0.026, HR 123.939, 95% CI could not be generated with statistical robustness due to  
 263 small sample size

264 Fig. 2C: p-value=0.029, HR 79.688, 95% CI could not be generated with statistical robustness due to  
 265 small sample size

266 Fig.2D: p-value=0.00048, HR and 95% CI could not be generated with statistical robustness due to  
 267 small sample size

268 HR: hazard ratio, CI: confidence interval

269 **Table 0. Impact of the overall and specific *PIK3CA* mutation status on relapse-free**  
 270 **survival (RFS)**

	p-value	Hazard Ratio Exp (B)	95% CI
Association between <i>PIK3CA</i> mutation & RFS	<b>**&lt;0.001</b>	<b>26.187</b>	<b>2.674 – 256.475</b>
	<b><i>PIK3CA</i> mutated</b>	<b><i>PIK3CA</i> not mutated</b>	
Mean RFS (months)	26.5	54.5	
95% CI of mean	13.99 - 39.01	51.12 - 57.84	



Table 3 continued

Median RFS in months (95% CI)	37.00 (0.00-81.91)			
<i>PIK3CA</i> Mutation	Relapse (n)	No relapse (n)	Total (n)	p-value
H1047R	2	2	4	
Other	1	1	2	
Total	3	3	6	0.281

\*\* : Association is significant at 1% level

CI: confidence interval, RFS: Relapse-free survival

“Other” includes the E542K mutation and the H1047R/E542K co-mutation

271

272 However, no significant association was found between individual *PIK3CA* hotspot mutations  
273 (H1047R or E542K) and RFS (Table 3).

## 274 Impact of selected clinicopathological features on RFS

275 LN metastasis was significantly associated with RFS (Table 4). Patients with LN-positive  
276 status had 123.94 times higher hazard of relapse compared to LN-negative patients (p=0.026,  
277 HR 123.94; Fig 2B). Additionally, a significant association was observed between Ki67  
278 proliferative status and RFS, with patients having higher proliferative indices facing a 79.69  
279 times higher hazard of relapse compared to those with Ki67 indices <20% (p=0.029, HR 79.69;  
280 Fig 2C).

281

## 282 Table 4. Univariate analysis of relapse-free survival (RFS) with selected 283 clinicopathological parameters

284

Parameter	Relapse (n)	No relapse (n)	Total	p-value
All	4	38	42	
<b>Age at diagnosis</b>				
≤50 years	1	10	11	
>50 years	3	28	31	0.772
<b>Histological type</b>				
IBC-NST	3	31	34	
Others	1	7	8	0.701
<b>Tumor grade/ differentiation</b>				
G1/ well-differentiated	1	13	14	

Table 4 continued

G2 or G3/ moderately or poorly differentiated	3	23	26	0.717
<b>Tumor size</b>				
≤ 2 cm	1	9	10	
> 2 cm	2	21	23	0.941
Mean (cm)	3.164			0.830
<b>Pathological tumor stage</b>				
Early stage (pT1/pT2)	3	26	29	
Late stage (pT3/pT4)	0	3	3	0.585
<b>Lymph node metastasis</b>				
Negative	0	19	19	
Positive	3	12	15	<b>*0.026</b>
<b>ER status</b>				
ER+	4	37	41	
ER-	0	1	1	0.769
<b>PR status</b>				
PR+	3	30	33	
PR-	1	8	9	0.805
<b>SHR status</b>				
SHR+	4	37	41	
SHR-	0	1	1	0.769
<b>HER2 status</b>				
HER2+	0	3	3	
HER2-	4	34	38	0.529
<b>Receptor based subtype</b>				
SHR+/HER2-	4	33	37	
Other	0	5	5	0.417
<b>Ki67 Proliferation index</b>				
<20%	0	19	19	
≥20%	4	15	19	<b>▲ 0.029</b>
<b>IHC-based molecular subtype</b>				
Luminal A/B	4	33	37	
HER2/TNBC	0	4	4	0.478

\*: association is significant at 5% level, HR 123.929, A statistically robust 95% CI not generated due to small sample number

▲: association is significant at 5% level, HR 79.688, A statistically robust 95% CI not generated due to small sample number

CI: confidence interval, ER: estrogen receptor, HER2: human epidermal growth factor receptor 2, HR: hazard ratio, IBC-NST: invasive breast carcinoma of no special type, IHC: immunohistochemical, LN: Lymph Node, PR: progesterone receptor, pT: pathological tumor stage, SHR: steroid hormone receptor, TNBC: triple-negative breast cancer

285

286

287

288 No other clinicopathological characteristics showed significant associations with RFS.

289 Finally, the multivariate Cox proportional hazard model (Fig 2D) demonstrated a significant  
290 association between the presence of a *PIK3CA* mutation, LN metastasis and a high Ki67 index  
291 with RFS ( $p < 0.001$ ).

292

## 293 Discussion

294 This study, to our knowledge, is the first to profile *PIK3CA* mutations, assess their prognostic  
295 value, and evaluate the implications of patient ancestry on somatic mutations in Sri Lankan  
296 breast cancer patients. Despite being a pilot study conducted in an underrepresented island  
297 population, our findings lay the groundwork for future research, including meta-analyses in  
298 larger cohorts.

299 The analysis of the clinicopathological parameters showed a median age of 58 years at breast  
300 cancer diagnosis, higher than the median age reported in a Sri Lankan study done over a decade  
301 ago, and higher than those reported from several Southeast Asian countries and Arab countries  
302 [26-29]. This could be due to Sri Lanka having one of the fastest ageing populations in the  
303 world [26]. The exponentially growing population of older women [26] would thus raise the  
304 incidence of breast cancer among this demographic, making it a key factor in the observed age  
305 shift in diagnosis. However, the fact that the majority of patients were already at stage 2 at the  
306 time of diagnosis, even with slow-growing cancers (i.e., luminal A and B), suggests a much  
307 earlier onset of the disease. This trend, consistent with other Asian countries, could probably  
308 be attributed to similar genetic and sociocultural factors and contrasts with developed Western  
309 countries, where median age of diagnosis occurs much later [30].

310 Similar to reports by previous studies conducted in Sri Lanka [31,32], it is noticeable that a  
311 majority of the tumors in our cohort were invasive breast carcinomas (IBC-NST) of tumor

312 grade 2 and/or stage II at diagnosis. This could be likely due to the absence of an established  
313 screening program aimed at early detection in Sri Lanka, thus reducing the likelihood of the  
314 cancer being detected at stage I and worsening disease outcomes [31]. These observations  
315 therefore highlight the need for the implementation of an effective screening system for breast  
316 cancer in Sri Lanka, to facilitate early diagnosis and ultimately reduce the overall burden of  
317 breast cancer in the country. Consistent with both local and global statistics [31-33], most  
318 tumors were ER/PR+ and HER2-, confirming that SHR+/HER2- luminal subtypes are the most  
319 common in Sri Lanka.

320 One of the study's primary objectives was to catalog *PIK3CA* mutations in the cohort of patients  
321 and to detect associations with clinicopathological features. The detection of H1047R and  
322 E542K mutations in 17.46% of cases falls within the prevalence range of *PIK3CA* mutations  
323 reported in many global studies on breast cancer [5-10]. Notably, this study also provides the  
324 first evidence of *PIK3CA* mutations in Sri Lankan breast cancer patients. India, the closest  
325 nation to Sri Lanka both geographically and genetically, reports a *PIK3CA* mutation prevalence  
326 of 23.2% [34], closely aligned with our findings. This similarity likely reflects common  
327 hereditary factors shaped by historical migrations and cultural exchanges, which may have  
328 influenced the prevalence of *PIK3CA* mutations in the South Asian region.

329 Although the predominance of H1047R as the most frequently observed *PIK3CA* mutation  
330 mirrors observations in other populations [35], our study reports a much a higher frequency of  
331 81.81% for H1047R among the total *PIK3CA* mutations, compared to the frequency ( $\approx$ 58%)  
332 reported in other studies [35,36]. A particularly rare and unexpected observation was the  
333 absence of the E545K mutation in our cohort, which is commonly reported as a moderately  
334 prevalent *PIK3CA* mutation ( $\approx$ 10%) in many other global populations [37,38]. This  
335 observation aligns with findings from an Iranian study, where the E545K mutation was

336 similarly absent [39]. Given that over 96% of our cohort comprises Sri Lankan Sinhalese and  
337 Sri Lankan Burghers (both of Indo-Aryan ancestry), this shared absence is likely to be  
338 stemming from a common ancestral link with Iranians, as both Iranians and Indo-Aryans are  
339 descendants of the Indo-Iranian branch of the Indo-European language family [23,24,40]. The  
340 absence of the E545K mutation, may thus be tied to ancestral roots, and consequent genetic  
341 backgrounds, among other possible factors. This provides valuable preliminary data from Sri  
342 Lanka, contributing to the global effort of cataloging *PIK3CA* mutations, which are highly  
343 heterogeneous across populations with different ancestral origins [10].

344 The significant association between patient ancestry and the presence of a *PIK3CA* mutation  
345 ( $p=0.028$ ), where the Indo-Aryan and Dravidian ancestries could potentially be associated with  
346 differing rates of *PIK3CA* mutations, suggests that ancestral influence tied to ethnic  
347 backgrounds could affect the occurrence of *PIK3CA* mutations in Sri Lanka as well as in South  
348 Asia. Though the strength of this observation is hampered by the small sample size, particularly  
349 with respect to the patients of Dravidian origin (one Tamil and one Moor in the entire cohort),  
350 the presence of *PIK3CA* mutations in both of them cannot be disregarded and warrant further  
351 ancestry-oriented mutational analyses. While this association aligns with global findings that  
352 somatic *PIK3CA* mutations are linked with ancestry across all cancer types including breast  
353 cancer [19-22], our study fills an important gap by investigating these ancestry-specific  
354 associations in a cohort, exclusively of South Asian ancestry. These insights may have  
355 significant implications for personalized treatment strategies, which may need to be tailored  
356 based on the patient's ancestry, as mutation profiles observed in one ancestry may not apply to  
357 others – in turn influencing treatment efficacy and outcomes.

358 Another notable observation was that the *PIK3CA* mutations in the cohort were exclusively  
359 seen in SHR+/HER2- luminal tumors, consistent with a similar subtype-specific dominance of

360 the mutation in previous studies [9]. The molecular mechanisms, including cross talk between  
361 the PI3K/AKT/mTOR and ER signaling pathways are postulated to be mediating this, with  
362 *PIK3CA* mutations driving oncogenesis via ER signaling, thereby predominating in SHR+  
363 breast cancer subtypes [17].

364 The significant association between *PIK3CA* mutations and increased likelihood of axillary LN  
365 metastasis ( $p=0.036$ , OR 9.60) highlights the potential prognostic value of *PIK3CA* mutations  
366 in predicting breast cancer progression in the Sri Lankan context. LN metastasis typically  
367 indicates a more advanced disease stage, a higher risk of distant metastases and worse  
368 outcomes [41]. Therefore, *PIK3CA* mutation testing could guide therapeutic strategies,  
369 particularly in decisions related to axillary management and risk assessment, with more  
370 intensive therapeutic approaches needed by patients with tumor-positive axillary LNs [42].  
371 Divergent findings in other studies [43], may reflect differences in genetic makeup and lifestyle  
372 factors, further emphasizing the need for nuanced molecular investigations in underrepresented  
373 populations like Sri Lanka.

374 The absence of associations between clinicopathological features and age group may indicate  
375 that other parameters, such as histological grade and Ki67 index, hold greater prognostic value  
376 than age, particularly in early-stage breast cancer cohorts [44], like ours. This is supported by  
377 the significant association between higher Ki67 indices and an increased risk of developing  
378 higher-grade tumors ( $p=0.024$ , OR 4.29). These findings further support the use of Ki67 index  
379 as a valuable prognostic marker [45] in Sri Lankan breast cancer patients and its importance in  
380 clinical decision-making.

381  
382 The striking association between *PIK3CA* mutations and reduced RFS in Sri Lankan breast  
383 cancer patients, with high statistical significance ( $p<0.001$ , HR 26.19, 95% CI 2.67 – 256.47)  
384 suggests that these mutations have a considerable impact on disease prognosis. The broad

385 confidence interval reflects the small sample size, but even the lower bound presents a  
386 significantly increased risk of relapse, nearly three times greater than patients without a  
387 *PIK3CA* mutation. This implies a significant impact of *PIK3CA* mutations on disease  
388 progression and brings out its potential clinical relevance as a prognostic biomarker in Sri  
389 Lankan breast cancer patients. As the first study to analyze *PIK3CA* mutations in Sri Lankan  
390 breast cancer patients, our findings are consistent with results from other Asian studies as well  
391 as studies on other malignancies, such as colorectal cancer [13,46,47].

392 However, our results did not show mutation-specific associations with RFS, which may be  
393 attributed to the smaller cohort size, necessitating further investigation with larger studies.  
394 Nevertheless, the predominance of the H1047R mutation- with an observed frequency of 0.81  
395 within the total mutations, suggests that it is likely to be responsible for reducing RFS [35] in  
396 our cohort, mirroring findings of the aforementioned Iranian study [39]. Given the shared Indo-  
397 Iranian ancestry of the Sinhalese and Iranians, this parallel may hint at ancestral genetic traits  
398 influencing both mutation occurrence and its clinical impact. This further emphasizes the need  
399 for nuanced investigations and personalized treatment approaches in populations with specific  
400 ancestral backgrounds.

401

402 Our analysis of clinicopathological characteristics also revealed that LN metastasis and Ki67  
403 proliferative index are key determinants of relapse risk. Patients with either tumor-positive LNs  
404 or a higher Ki67 index ( $\geq 20\%$ ), exhibited a significantly greater hazard of relapse compared to  
405 their counterparts without tumor deposits in LNs ( $p=0.026$ , HR 123.93) and with lower  
406 proliferative indices ( $p=0.029$ , HR 79.68). These results reinforce the established role of LN  
407 metastasis and cellular proliferation rates as reliable indicators of relapse and survival in breast  
408 cancer patients [48-50], while also suggesting underlying molecular mechanisms that

409 contribute to cancer recurrence [50]. This further underscores the importance of these factors  
410 in staging, prognosis, and therapy of invasive breast cancer in Sri Lankan patients.

411  
412 One of the most paramount points in our study is the validation of the observed association  
413 between LN metastasis and RFS ( $p=0.026$ ) by the compounded association of *PIK3CA*  
414 mutations with both LN metastasis ( $p=0.036$ ) and significantly reduced RFS ( $p<0.001$ ). This  
415 enables us to bridge all three findings. Additionally, the multivariate model demonstrated the  
416 compounded impact of all these factors ( $p<0.001$ ) on the likelihood of disease relapse,  
417 ultimately emphasizing the importance of considering all three factors- *PIK3CA* mutations, LN  
418 status, and Ki67 index- along with patient ancestry, in prognosis and treatment planning for  
419 breast cancer patients in Sri Lanka and across South Asia. This may also suggest an underlying  
420 mechanistic link that contributes to the aggressive behavior of breast cancer. It is plausible that  
421 in ancestries more prone to the occurrence of *PIK3CA* mutations in malignant cells, these  
422 mutations promote the cancer spread via lymphatic channels, leading to LN metastasis and  
423 subsequent recurrence. This progression highlights a crucial step in cancer spread, as LNs  
424 function as critical filtering sites within the lymphatic system, where cancer cells infiltrate and  
425 establish secondary tumor deposits [52].

426 The above findings align with Stephen Paget's 'Seed and soil hypothesis,' which proposes that  
427 metastasis occurs when certain tumor cells (the "seed") preferentially grow in particular organ  
428 microenvironments (the "soil") - and that metastases only develop when the right seed is  
429 implanted in the right soil [53]. Our study suggests that *PIK3CA* driver mutations, possibly  
430 the predominant H1047R mutation, may act as a genetic change that enhances the metastatic  
431 potential of breast cancer cells by promoting survival and proliferation. These genetic  
432 alterations could facilitate the invasion of lymphatic vessels, leading to the establishment of  
433 metastatic colonies in LNs. In this context, *PIK3CA* mutations may assist in planting the "seed"



434 in the "soil" by promoting the establishment of tumor deposits in axillary LNs. Strikingly,  
435 further developments to this hypothesis, elaborate how the above mechanisms could result in  
436 the primary breast cancer metastasizing to the lungs and bones [54], the exact two organs of  
437 relapse diagnosed in our *PIK3CA*-mutated patient population. Therefore, by elucidating this  
438 molecular pathway, our study offers valuable insights into the potential role of *PIK3CA* in the  
439 underlying biology of aggressive breast cancers, particularly in the Sri Lankan population.

440

441 To our knowledge, PI3K inhibitors are not currently prescribed for breast cancer patients in Sri  
442 Lanka. Our study's preliminary observation of a negative prognostic outcome in the  
443 predominantly SHR+/HER2-, Sri Lankan breast cancer cohort with *PIK3CA* mutations, serves  
444 as a calling out for Sri Lankan clinicians to incorporate *PIK3CA* mutational testing in routine  
445 diagnostic protocols. Patients identified with these mutations could benefit from PI3K $\alpha$   
446 inhibitors such as Alpelisib, which are approved for this clinical scenario and are known to  
447 prolong the RFS in patients [55]. Furthermore, this treatment approach could be tailored for  
448 high risk ethnic groups within Sri Lanka (i.e. those with Dravidian ancestry), should a larger  
449 study with a better representation of ethnic minority groups, confirm the association of *PIK3CA*  
450 mutations with patient ancestry. This personalized approach could extend beyond breast cancer  
451 to other malignancies such as colorectal, gastric, ovarian, and cervical cancers, which have also  
452 exhibited *PIK3CA* mutations and consequent impacts on prognosis [8,56].

453

454 Despite the significant implications of our study, there are several limitations. The primary  
455 limitation is the sample size, which was restricted due to financial and logistical constraints.  
456 This limitation impacted the statistical power of several results and affected the ability to  
457 perform robust regression and multivariate analyses. Additionally, the retrospective nature of  
458 our study also introduced certain constraints. For example, the grouping of patients into broad

459 ancestral categories was based on self-reported ethnicity in hospital records of the patients,  
460 rather than genetic testing. Although studies indicate a high concordance between self-reported  
461 ancestry and genetic testing results [22], potential misclassification may have oversimplified  
462 the complex genetic landscape of the Sri Lankan population. Furthermore, the  
463 underrepresentation of minority ethnic groups due to the uneven geographical distribution of  
464 these populations, greatly impaired the power of the detected association of ancestry with  
465 mutation status. Overcoming this limitation might require recruiting patients from a broader  
466 array of hospitals and oncology clinics around the country, to ensure adequate representation  
467 of all ancestral groups.

468 Another limitation is the absence of data on other potentially interacting mutations within the  
469 *PIK3CA* gene or in other OGs and TSGs that were not tested in the cohort, and their interplay  
470 with the tested mutations. Additionally, the follow-up period in our study was relatively short.  
471 All patients with survival data (n=42) were alive, and 90.5% remained recurrence-free within  
472 5 years post-diagnosis, which limited the survival analysis to RFS as a proxy. A longer follow-  
473 up period would provide a more comprehensive understanding of long-term outcomes.  
474 Nevertheless, this pilot study, the first to analyze *PIK3CA* mutations in relation to breast cancer  
475 in different ancestries in Sri Lanka, together with its observed prognostic implications, offers  
476 sufficient insight and direction to the chosen objectives, so that a larger, more conclusive study  
477 can be conducted to confirm and validate the reported findings. Accordingly, a study on a larger  
478 Sri Lankan cohort with a subsequent functional analysis is necessary to address the above  
479 limitations, especially to confirm the absence of the E545K mutation in Sri Lankan breast  
480 cancer patients and to validate the significance of observed mutations, particularly in the  
481 Dravidian population, which appears to be at higher risk based on our findings.

## 482 **Conclusion**

483 This study is the first to report on *PIK3CA* mutations in Sri Lankan breast cancer patients, a  
484 population with a unique genetic heritage, demonstrating a *PIK3CA* mutation prevalence of  
485 17.46% and significant associations between ancestry-linked mutations, LN metastasis and  
486 cancer relapse. Our findings also suggest the existence of a possible molecular link contributing  
487 to metastasis.

488 Notably, our observation of a significant association between patient ancestry and *PIK3CA*  
489 mutation status is the first of its kind in an exclusively South Asian cohort. This invites  
490 comparisons with similar findings - such as the absence of the E545K mutation - in global  
491 populations with shared ancestries. Our findings also provide a foundation for prognostic  
492 stratification and therapeutic decision-making for breast cancer patients in Sri Lanka - rooted  
493 in the presence of *PIK3CA* mutations and patient ancestry - for identifying high-risk patient  
494 subsets who may benefit from targeted therapies such as PI3K inhibitors, which are not  
495 currently prescribed in Sri Lanka. Further large-scale functional studies are necessary to  
496 validate these findings and to improve clinical outcomes for breast cancer patients in Sri Lanka.

497

## 498 **Acknowledgements:**

499 We gratefully acknowledge the invaluable support provided by the staff at the Centre for  
500 Immunology and Molecular Biology (CIMB) and the Centre for Transgenic Technologies at  
501 the Faculty of Science, University of Colombo, Sri Lanka, as well as the medical laboratory  
502 technicians at the Diagnostics Laboratory of University Hospital KDU, Sri Lanka in the  
503 completion of the laboratory component of this research.

## 504 **References:**

- 505 1. Barakeh DH, Aljelaify R, Bashawri Y, Almutairi A, Alqubaishi F, Alnamnakani M, et  
506 al. Landscape of somatic mutations in breast cancer: new opportunities for targeted  
507 therapies in Saudi Arabian patients. *Oncotarget* [Internet]. 2021 Mar 30 [cited 2022  
508 Mar 24];12(7):686–97. Available from:  
509 <https://www.oncotarget.com/article/27909/text/>
- 510 2. Fan L, Goss PE, Strasser-Weippl K. Current Status and Future Projections of Breast  
511 Cancer in Asia. *Breast Care*. 2015;10(6):372–8.
- 512 3. Zhukova LG, Andreeva II, Zavalishina LE, Zakiriakhodzhaev AD, Koroleva IA,  
513 Nazarenko AV, et al. Breast cancer. Современная онкология [Internet]. 2021 May 1  
514 [cited 2023 Jan 16];23(1):5–40. Available from:  
515 <https://doaj.org/article/30b4e9318d3543faa6dfbfe39ff5c88d>
- 516 4. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature*. 2009  
517 Apr;458(7239):719–24.
- 518 5. Miricescu D, Totan A, Stanescu-Spinu II, Badoiu SC, Stefani C, Greabu M.  
519 PI3K/AKT/mTOR Signaling Pathway in Breast Cancer: From Molecular Landscape  
520 to Clinical Aspects. *International Journal of Molecular Sciences*. 2020 Dec  
521 26;22(1):173.
- 522 6. Bachman KE, Argani P, Samuels Y, Silliman N, Ptak J, Szabo S, et al. The PIK3CA  
523 gene is mutated with high frequency in human breast cancers. *Cancer Biology &  
524 Therapy*. 2004 Aug;3(8):772–5.
- 525 7. Samuels Y, Waldman T. Oncogenic Mutations of PIK3CA in Human Cancers.  
526 *Current Topics in Microbiology and Immunology* [Internet]. 2010 [cited 2019 Oct  
527 8];21–41. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3164550/>

- 528 8. Karakas B, Bachman KE, Park BH. Mutation of the PIK3CA oncogene in human  
529 cancers. *British Journal of Cancer* [Internet]. 2006 Jan 31 [cited 2019 Oct  
530 8];94(4):455–9. Available from: <https://www.nature.com/articles/6602970>
- 531 9. Reinhardt K, K Stückrath, Hartung C, Kaufhold S, Christoph Uleer, Hanf V, et al.  
532 PIK3CA-mutations in breast cancer. *Breast cancer research and treatment* [Internet].  
533 2022 Oct 24;196(3):483–93. Available from:  
534 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9633529/>
- 535 10. Chen JW, Murugesan K, Newberg JY, Sokol ES, Savage HM, Stout TJ, et al.  
536 Comparison of PIK3CA Mutation Prevalence in Breast Cancer Across Predicted  
537 Ancestry Populations. *JCO Precision Oncology*. 2022 Nov;(6).
- 538 11. Li SY, Rong M, Grieu F, Iacopetta B. PIK3CA mutations in breast cancer are  
539 associated with poor outcome. *Breast Cancer Research and Treatment*. 2005 Nov  
540 30;96(1):91–5.
- 541 12. Hu H, Zhu J, Zhong Y, Geng R, Ji Y, Guan Q, et al. PIK3CA mutation confers  
542 resistance to chemotherapy in triple-negative breast cancer by inhibiting apoptosis  
543 and activating the PI3K/AKT/mTOR signaling pathway. *Annals of Translational  
544 Medicine* [Internet]. 2021 Mar 1 [cited 2021 Dec 21];9(5):410–0. Available from:  
545 <https://atm.amegroups.com/article/view/64371/html>
- 546 13. Cho YA, Ko SY, Suh YJ, Kim S, Park JH, Park HR, et al. PIK3CA Mutation as  
547 Potential Poor Prognostic Marker in Asian Female Breast Cancer Patients Who  
548 Received Adjuvant Chemotherapy. *Current Oncology*. 2022 Apr 19;29(5):2895–908.
- 549 14. Fillbrunn M, Signorovitch J, André F, Wang I, Lorenzo I, Ridolfi A, et al. PIK3CA  
550 mutation status, progression and survival in advanced HR + /HER2- breast cancer: a  
551 meta-analysis of published clinical trials. *BMC Cancer*. 2022 Sep 21;22(1).

- 552 15. Zagami P, Aranzazu Fernandez-Martinez, Rashid NU, Hoadley KA, Spears PA,  
553 Curigliano G, et al. Association of PIK3CA Mutation With Pathologic Complete  
554 Response and Outcome by Hormone Receptor Status and Intrinsic Subtype in Early-  
555 Stage ERBB2/HER2-Positive Breast Cancer. *JAMA Network Open*. 2023 Dec  
556 20;6(12):e2348814–4.
- 557 16. Cizkova M, Susini A, Vacher S, Cizeron-Clairac G, Andrieu C, Driouch K, et al.  
558 PIK3CA mutation impact on survival in breast cancer patients and in ER $\alpha$ , PR and  
559 ERBB2-based subgroups. *Breast Cancer Research : BCR* [Internet]. 2012;14(1):R28.  
560 Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3496146/>
- 561 17. Zardavas D, te Marvelde L, Milne RL, Fumagalli D, Fountzilias G, Kotoula V, et al.  
562 Tumor PIK3CA Genotype and Prognosis in Early-Stage Breast Cancer: A Pooled  
563 Analysis of Individual Patient Data. *Journal of Clinical Oncology*. 2018 Apr  
564 1;36(10):981–90.
- 565 18. Wang S, Pitt JJ, Zheng Y, Yoshimatsu TF, Gao G, Sanni A, et al. Germline variants  
566 and somatic mutation signatures of breast cancer across populations of African and  
567 European ancestry in the US and Nigeria. *International Journal of Cancer*. 2019 Jun  
568 7;145(12):3321–33.
- 569 19. Pitt JJ, Riester M, Zheng Y, Yoshimatsu TF, Sanni A, Oluwasola O, et al.  
570 Characterization of Nigerian breast cancer reveals prevalent homologous  
571 recombination deficiency and aggressive molecular features. *Nature Communications*  
572 [Internet]. 2018 Oct 16 [cited 2021 Sep 22];9(1):4181. Available from:  
573 <https://www.nature.com/articles/s41467-018-06616-0#Sec10>
- 574 20. Keenan T, Moy B, Mroz EA, Ross K, Niemierko A, Rocco JW, et al. Comparison of  
575 the Genomic Landscape Between Primary Breast Cancer in African American Versus

- 576 White Women and the Association of Racial Differences With Tumor Recurrence.  
577 Journal of Clinical Oncology. 2015 Nov 1;33(31):3621–7.
- 578 21. Carrot-Zhang J, Chambwe N, Damrauer JS, Knijnenburg TA, Robertson AG, Yau C,  
579 et al. Comprehensive analysis of genetic ancestry and its molecular correlates in  
580 cancer. *Cancer Cell*. 2020 May;37(5). doi:10.1016/j.ccell.2020.04.012
- 581 22. Arora K, Tran TN, Kemel Y, Miika Mehine, Liu YL, Subhiksha Nandakumar, et al.  
582 Genetic Ancestry Correlates with Somatic Differences in a Real-World Clinical  
583 Cancer Sequencing Cohort. *Cancer Discovery*. 2022 Sep 1;12(11):2552–65.
- 584 23. Perera N, Galhena G, Ranawaka G. X-chromosomal STR based genetic  
585 polymorphisms and demographic history of Sri Lankan ethnicities and their  
586 relationship with global populations. *Scientific Reports*. 2021 Jun 17;11(1).
- 587 24. Singh PP., Kumar S, Pasupuleti N, Weerasooriya PR, van Driem G, Tennekoon KH,  
588 et al. Reconstructing the population history of Sinhalese, the major ethnic group in Śrī  
589 Lāṅkā. *iScience* [Internet]. 2023 Oct 1;26(10):107797–7. Available from:  
590 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10514440/>
- 591 25. Vithana PVSC, Dheerasinghe DSAF, Handagiripathira HMI, Alahapperuma S,  
592 Nilaweera I, Perera S. Sri Lankan Patterns of Female Cancers: Incidence and  
593 Mortality Over 1995-2010. *Asian Pacific Journal of Cancer Care*. 2021 Mar  
594 26;6(1):27–33.
- 595 26. Fernando A, Jayarajah U, Prabashani S, Fernando EA, Seneviratne SA. Incidence  
596 trends and patterns of breast cancer in Sri Lanka: an analysis of the national cancer  
597 database. *BMC Cancer*. 2018 Apr 27;18(1).
- 598 27. Anwar SL, Raharjo CA, Herviastuti R, Dwianingsih EK, Setyoheriyanto D, Avanti  
599 WS, et al. Pathological profiles and clinical management challenges of breast cancer

- 600 emerging in young women in Indonesia: a hospital-based study. *BMC Women's*  
601 *Health*. 2019 Feb 6;19(1).
- 602 28. Najjar H, Easson A. Age at diagnosis of breast cancer in Arab nations. *International*  
603 *Journal of Surgery*. 2010;8(6):448–52. doi:10.1016/j.ijss.2010.05.012
- 604 29. Abdel-Razeq H, Almasri H, Abdel Rahman F, Abdulelah H, Abu Nasser M, Salam  
605 M, et al. Clinicopathological Characteristics And Treatment Outcomes Of Breast  
606 Cancer Among Adolescents And Young Adults In A Developing Country. *Cancer*  
607 *Management and Research*. 2019 Nov 1;Volume 11:9891–7.
- 608 30. Leong SPL, Shen ZZ, Liu TJ, Agarwal G, Tajima T, Paik NS, et al. Is Breast Cancer  
609 the Same Disease in Asian and Western Countries? *World Journal of Surgery*  
610 [Internet]. 2010 [cited 2021 May 28];34(10):2308–24. Available from:  
611 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2936680/>
- 612 31. Balawardena J, Skandarajah T, Rathnayake W, Joseph N. Breast Cancer Survival in  
613 Sri Lanka. *JCO Global Oncology*. 2020 Apr;(6):589–99.
- 614 32. Perera R, Senanayake A, Balawardana J, Malviarachchi S, Kumarasinghe I, De Silva  
615 A. Clinicopathological analysis of breast cancers in a single unit of a Sri Lankan  
616 tertiary care hospital. *Sri Lanka Journal of Cancer* [Internet]. 2023;3(1):27–35.  
617 Available from: <https://slco.lk/wp-content/uploads/2023/09/j5.pdf>
- 618 33. Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LAG, et al. US  
619 Incidence of Breast Cancer Subtypes Defined by Joint Hormone Receptor and HER2  
620 Status. *JNCI: Journal of the National Cancer Institute* [Internet]. 2014 Apr 28;106(5).  
621 Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4580552/>
- 622 34. Ahmad F, Badwe A, Verma G, Bhatia S, Das BR. Molecular evaluation of PIK3CA  
623 gene mutation in breast cancer: determination of frequency, distribution pattern and



- 624 its association with clinicopathological findings in Indian patients. *Medical Oncology*.  
625 2016 Jun 9;33(7).
- 626 35. Mangone F, Bobrovnitchaia I, Salaorni S, Manuli E, Nagai M. PIK3CA exon 20  
627 mutations are associated with poor prognosis in breast cancer patients. *Clinics*. 2012  
628 Nov 7;67(11):1285–90.
- 629 36. Lai YL, Mau BL, Cheng WH, Chen HM, Chiu HH, Tzen CY. PIK3CA exon 20  
630 mutation is independently associated with a poor prognosis in breast cancer patients.  
631 *Annals of Surgical Oncology* [Internet]. 2008 Apr 1 [cited 2020 Nov 24];15(4):1064–  
632 9. Available from: <https://pubmed.ncbi.nlm.nih.gov/18183466/>
- 633 37. Martínez-Sáez O, Chic N, Pascual T, Adamo B, Vidal M, González-Farré B, et al.  
634 Frequency and spectrum of PIK3CA somatic mutations in breast cancer. *Breast*  
635 *Cancer Research*. 2020 May 13;22(1).
- 636 38. Chao T, Tsai Y, Liu C, Lien PJ, Lin Y, Feng CJ, et al. Prevalence of PIK3CA  
637 mutations in Taiwanese patients with breast cancer: a retrospective next-generation  
638 sequencing database analysis. *Frontiers in Oncology* [Internet]. 2023 Aug 15 [cited  
639 2024 Jan 24];13. Available from:  
640 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10466395/>
- 641 39. Nazeri E, Parniani M, Olfatbakhsh A, Najafi S, Peyman M, Karami B, et al.  
642 Evaluation of PIK3CA Gene Mutations in Breast Cancer Patients Treated by  
643 Trastuzumab. *Archives of Breast Cancer*. 2023 Jun 16;248–55.
- 644 40. Indo-Iranian languages [Internet]. *Encyclopedia Britannica*. Available from:  
645 <https://www.britannica.com/topic/Indo-Iranian-languages>
- 646 41. Luo M, Lin X, Hao D, Kangle Wang Shen, Wu W, Wang L, et al. Incidence and risk  
647 factors of lymph node metastasis in breast cancer patients without preoperative  
648 chemoradiotherapy and neoadjuvant therapy: analysis of SEER data. *Gland Surgery*

- 649 [Internet]. 2023 Nov 1 [cited 2024 Aug 27];12(11):1508–24. Available from:  
650 <https://pubmed.ncbi.nlm.nih.gov/38107495/>
- 651 42. Rao R. The Evolution of Axillary Staging in Breast Cancer. PubMed. 2015 Nov  
652 27;112(5):385–8.
- 653 43. Kalinsky K, Jacks LM, Heguy A, Patil S, Drobnjak M, Bhanot UK, et al. PIK3CA  
654 Mutation Associates with Improved Outcome in Breast Cancer. Clinical Cancer  
655 Research. 2009 Aug 11;15(16):5049–59.
- 656 44. Arikan AE, Kara H, Dülgeroğlu O, Erdoğan EN, Capkinoglu E, Uras C. Do  
657 prognosis and clinicopathological features differ in young early-stage breast cancer?  
658 Frontiers in Surgery [Internet]. 2022 Oct 19 [cited 2024 Sep 29];9. Available from:  
659 <https://doi.org/10.3389/fsurg.2022.900363>
- 660 45. Kanyılmaz G, Yavuz BB, Aktan A, Karaağaç M, Uyar M, Fındık S. Prognostic  
661 Importance of Ki-67 in Breast Cancer and Its Relationship with Other Prognostic  
662 Factors - European Journal of Breast Health [Internet]. Doi.org. 2019 [cited 2024 Sep  
663 29]. Available from: <https://doi.org/10.5152/ejbh.2019.4778>
- 664 46. Deng L, Zhu X, Sun Y, Wang J, Zhong X, Li J, et al. Prevalence and Prognostic Role  
665 of PIK3CA/AKT1 Mutations in Chinese Breast Cancer Patients. Cancer Research and  
666 Treatment. 2019 Jan 15;51(1):128–40.
- 667 47. Kato S, Iida S, Higuchi T, Ishikawa T, Takagi Y, Yasuno M, et al. PIK3CA mutation  
668 is predictive of poor survival in patients with colorectal cancer. International Journal  
669 of Cancer. 2007;121(8):1771–8.
- 670 48. Rakha EA, Tan PH, Varga Z, Tse GM, Shaaban AM, Climent F, et al. Prognostic  
671 factors in metaplastic carcinoma of the breast: a multi-institutional study. British  
672 Journal of Cancer. 2014 Nov 25;112(2):283–9.

- 673 49. Sun X, Kaufman PD. Ki-67: more than a proliferation marker. *Chromosoma*  
674 [Internet]. 2018 Jun 1;127(2):175–86. Available from:  
675 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5945335/>
- 676 50. Liu N, Yang Z, Liu X, Niu Y. Lymph node status in different molecular subtype of  
677 breast cancer: triple negative tumours are more likely lymph node negative.  
678 *Oncotarget*. 2017 Feb 2;8(33).
- 679 51. Hao Y, Baker D, ten Dijke P. TGF- $\beta$ -Mediated Epithelial-Mesenchymal Transition  
680 and Cancer Metastasis. *International Journal of Molecular Sciences* [Internet]. 2019  
681 Jun 5;20(11):2767. Available from:  
682 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6600375/>
- 683 52. Ji H, Hu C, Yang X, Liu Y, Ji G, Ge S, et al. Lymph node metastasis in cancer  
684 progression: molecular mechanisms, clinical significance and therapeutic  
685 interventions. *Signal Transduction and Targeted Therapy* [Internet]. 2023 Sep 27  
686 [cited 2023 Sep 29];8(1):1–33. Available from:  
687 <https://www.nature.com/articles/s41392-023-01576-4>
- 688 53. Paget S. THE DISTRIBUTION OF SECONDARY GROWTHS IN CANCER OF  
689 THE BREAST. *The Lancet*. 1889 Mar;133(3421):571–3.
- 690 54. Langley RR, Fidler IJ. The seed and soil hypothesis revisited-The role of tumor-  
691 stroma interactions in metastasis to different organs. *International Journal of Cancer*.  
692 2011 Mar 25;128(11):2527–35.
- 693 55. André F, Ciruelos E, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib  
694 for PIK3CA-Mutated, Hormone Receptor–Positive Advanced Breast Cancer. *New*  
695 *England Journal of Medicine*. 2019 May 16;380(20):1929–40.

696 56. Ogino S, Nosho K, Kirkner GJ, Shima K, Irahara N, Kure S, et al. PIK3CA Mutation  
697 Is Associated With Poor Prognosis Among Patients With Curatively Resected Colon  
698 Cancer. *Journal of Clinical Oncology*. 2009 Mar 20;27(9):1477–84.

699

## 700 **Supporting Information**

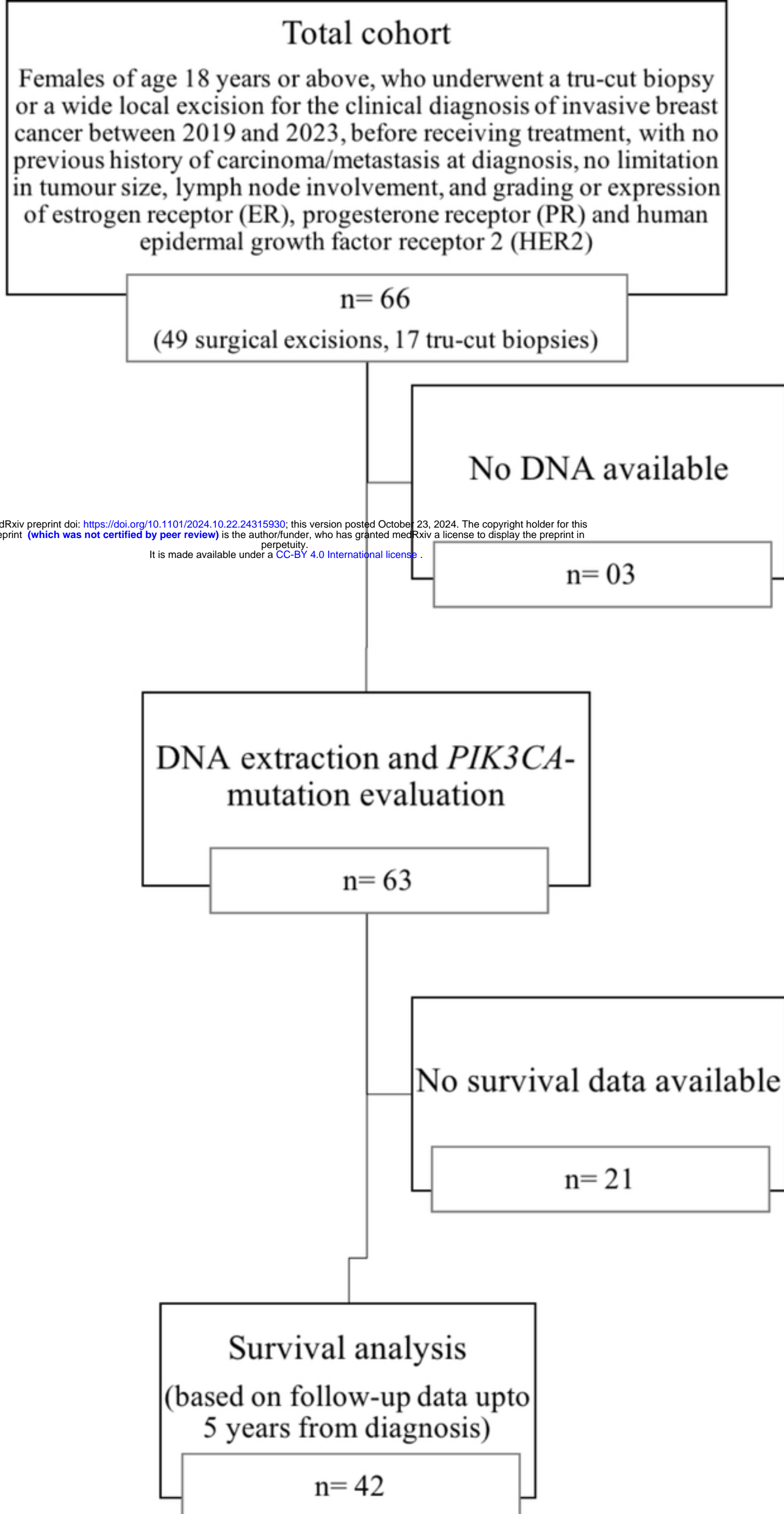
701 **S1 Fig. Ethical approval for the study granted by the Ethics Review Committee of the**  
702 **Institute of Biology, Sri Lanka (ERC IOBSL 303 07 2023)**

703 **S2 Fig. Age distribution of the cohort**

704 **S3 Fig. Distribution of tumor size (cm) in the cohort**

705 **S1 Table. Grouping of the variables used for the data analysis**

706 **S2 Table. Status of lymph node (LN) metastasis and patient age group in selected clinical**  
707 **and histopathological groups**



medRxiv preprint doi: <https://doi.org/10.1101/2024.10.22.24315930>; this version posted October 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Figure 1

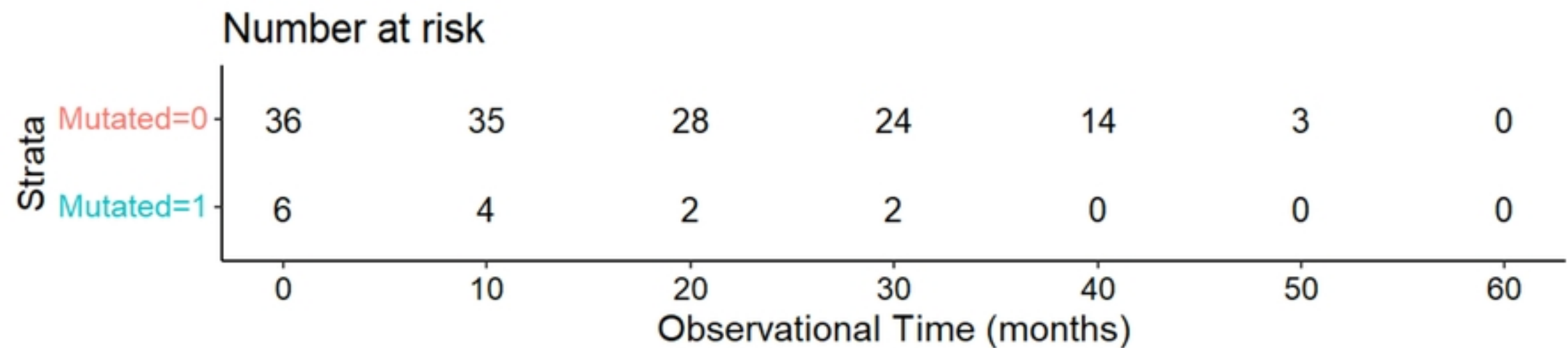
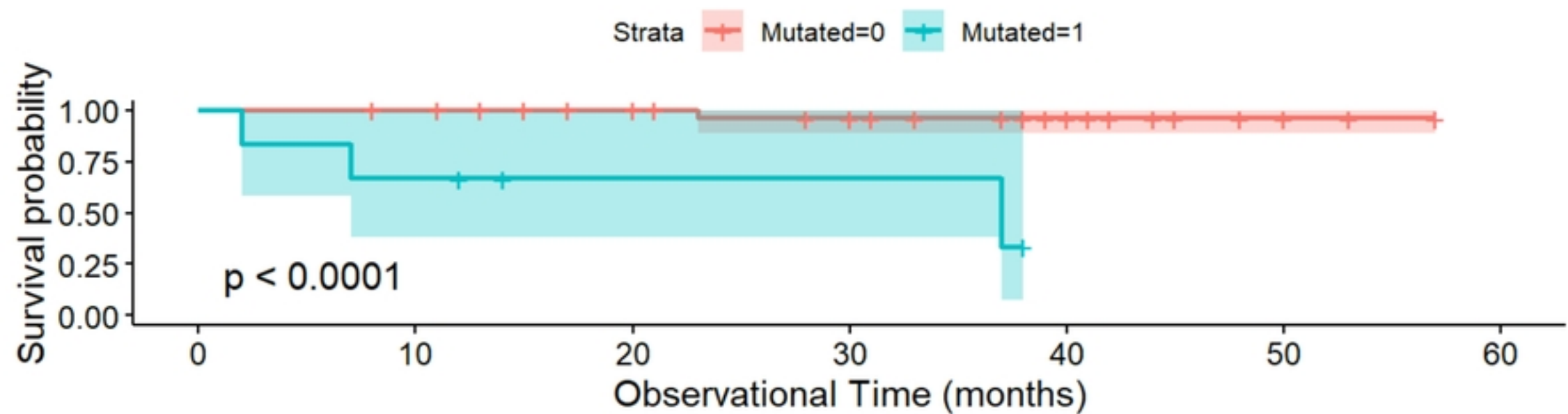


Figure 2 (A, B, C and D)

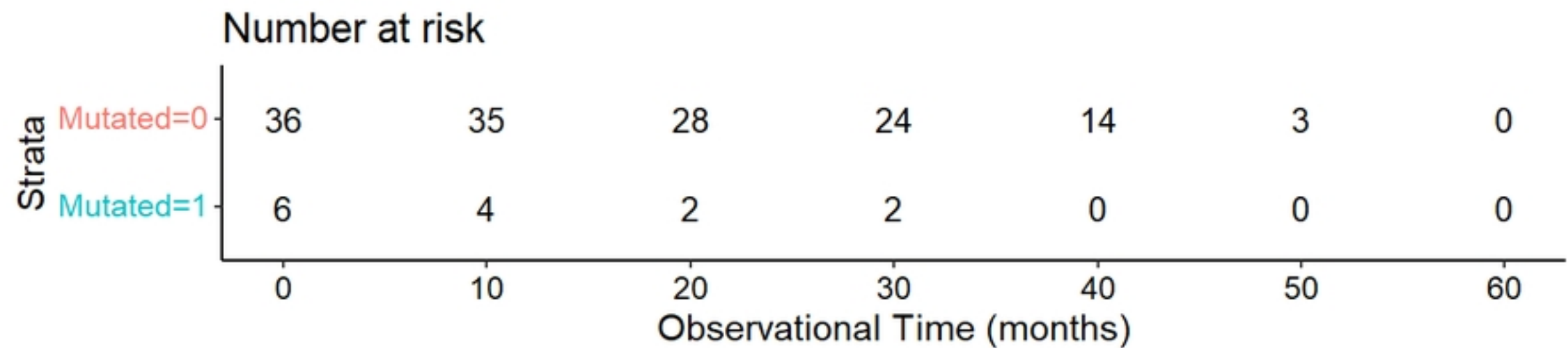
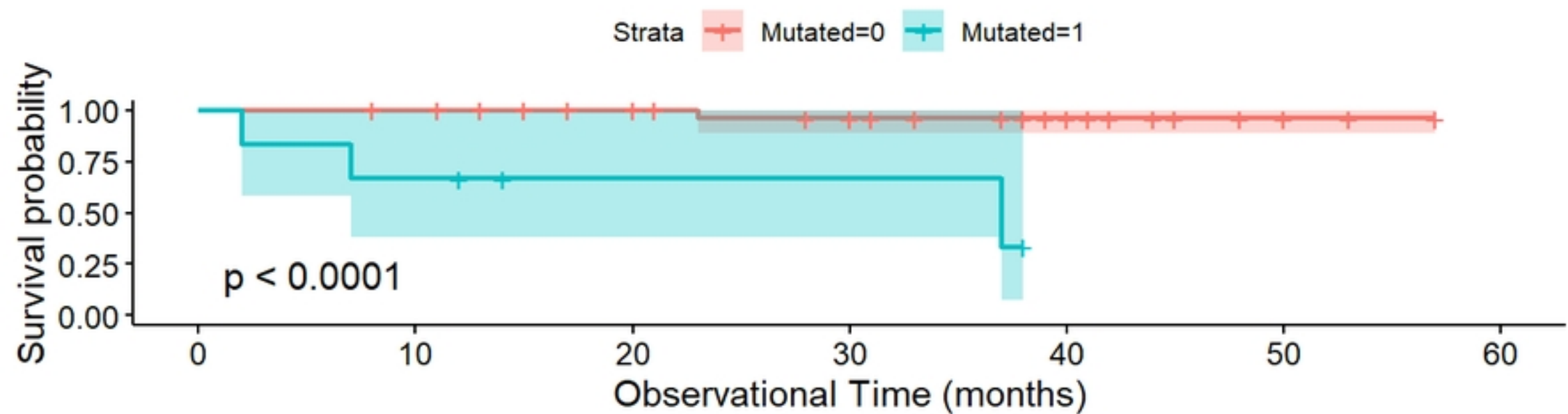


Figure 2A

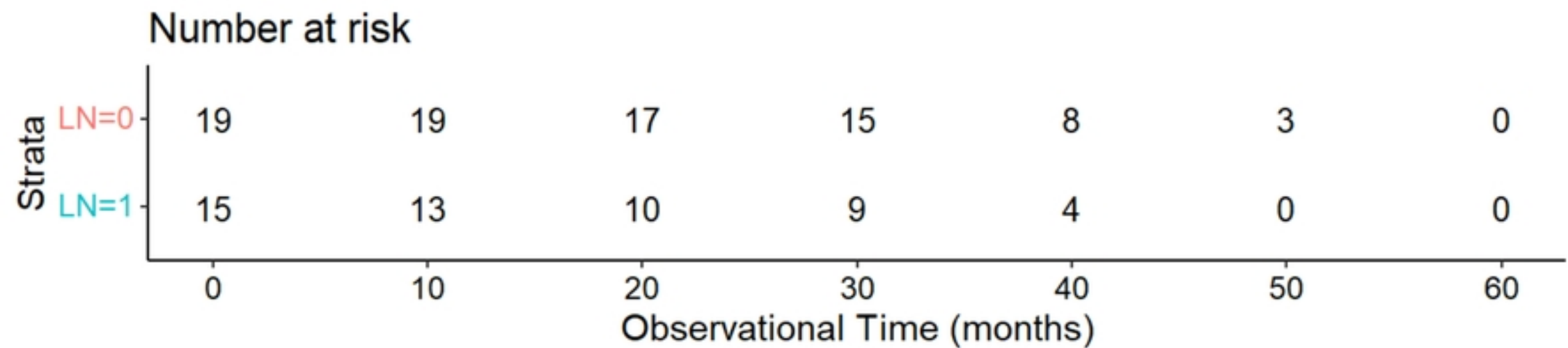
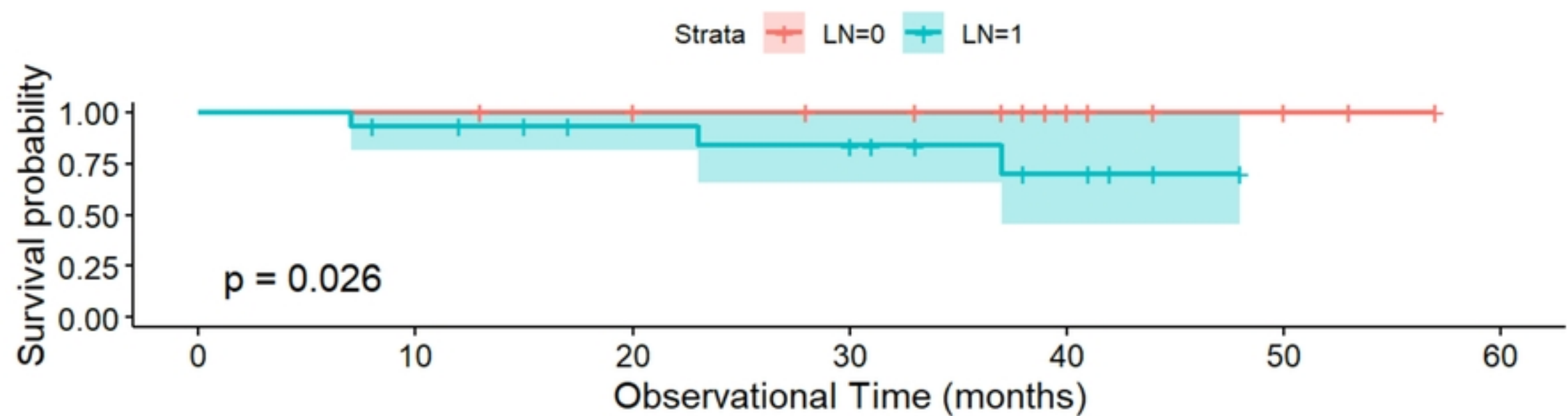


Figure 2B



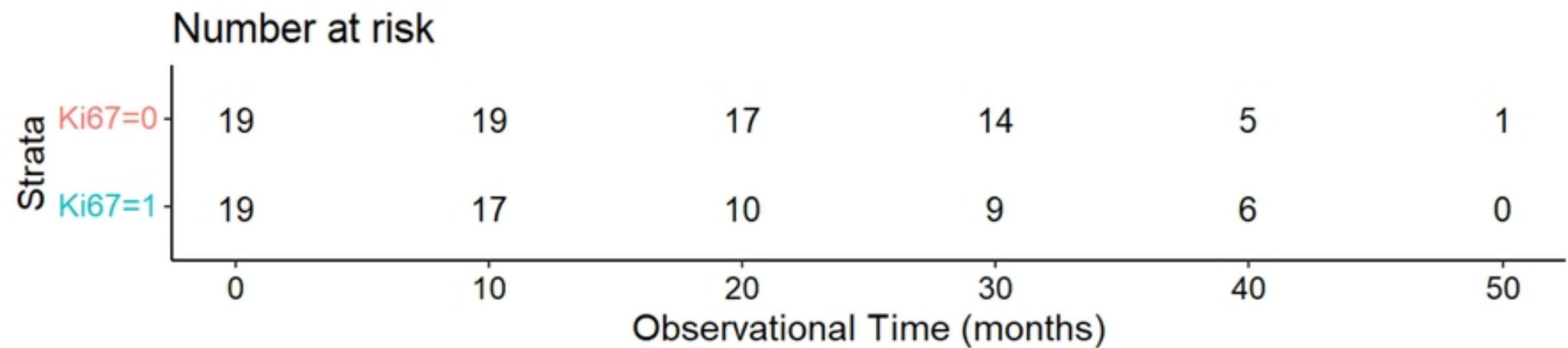
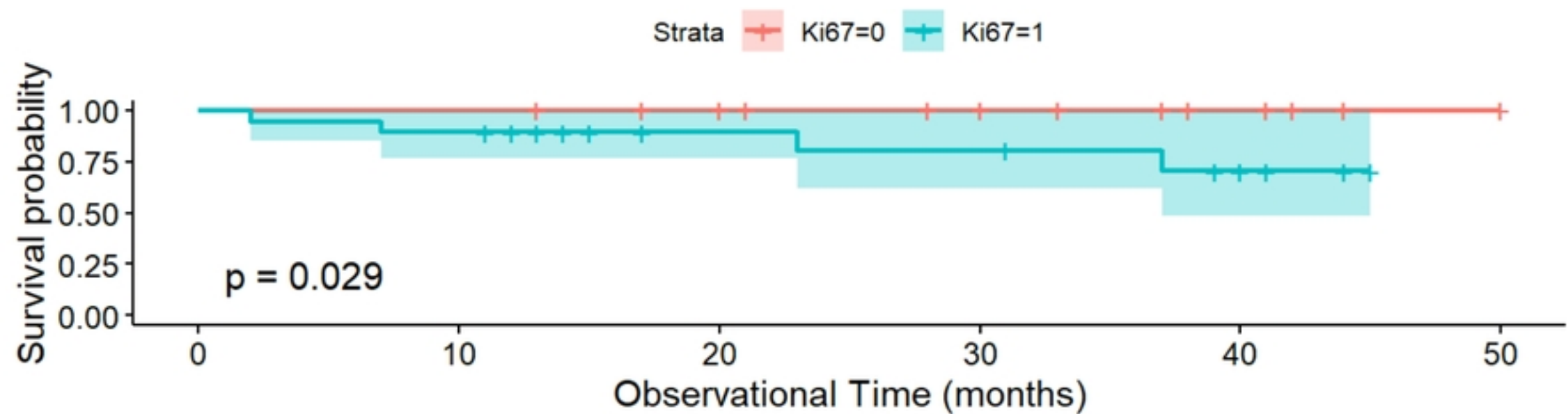


Figure 2C

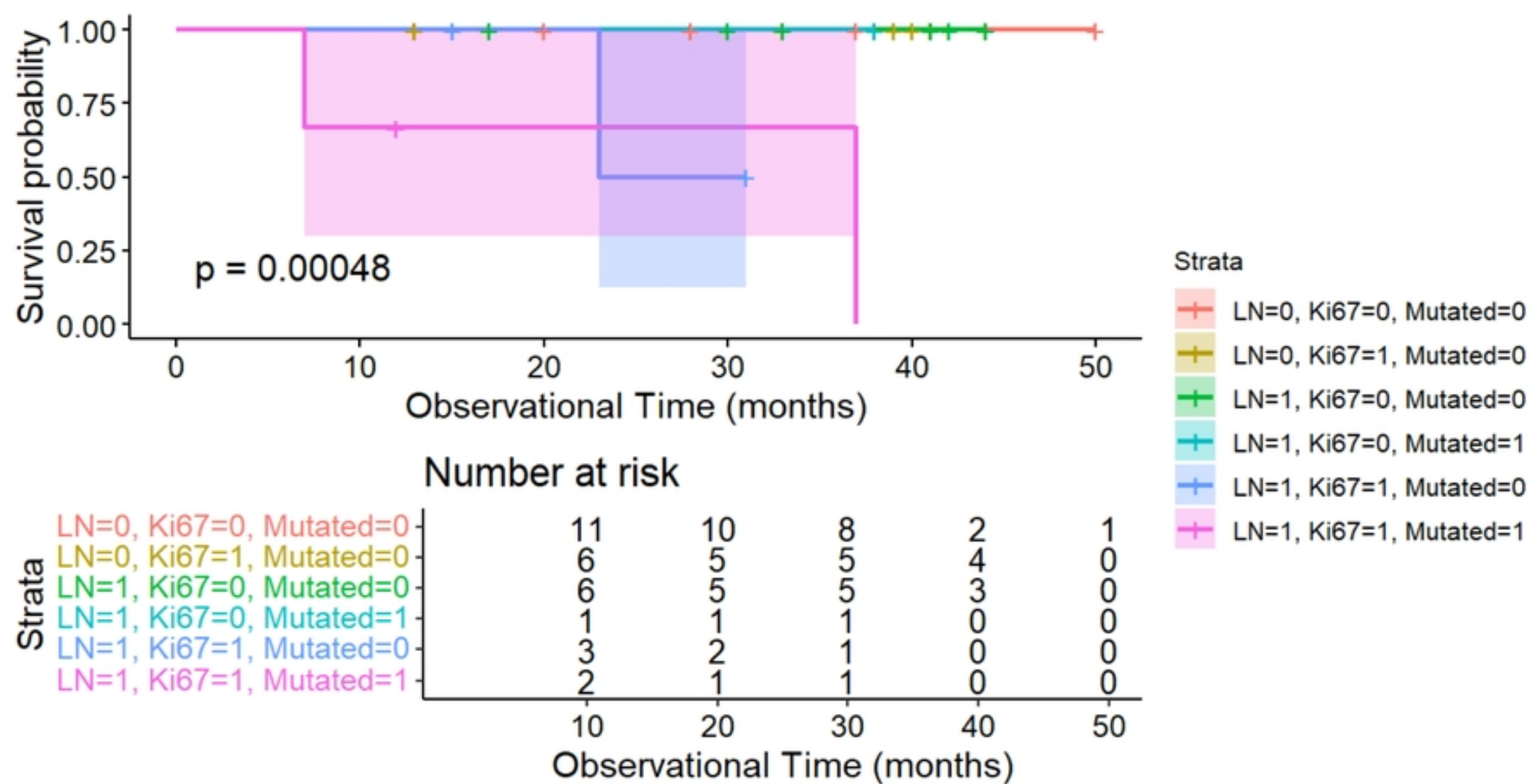


Figure 2D