

**Single Nucleotide Polymorphisms in the Base
Excision Repair and Double Strand Break Repair
pathway genes in Lung Cancer patients and their
correlation with clinical outcomes**

A Thesis

submitted in partial fulfillment of the requirements for the award of the degree of

Doctor of Philosophy

in

Department of Biotechnology

by

Amrita Singh

(Reg no: 901300004)

under the supervision of

Dr. Siddharth Sharma
Associate Professor

Dr. Navneet Singh
Additional Professor

Dr. D. Behera
Professor



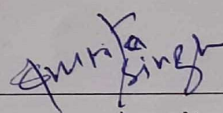
Thapar Institute of Engineering & Techonology
Patiala, India
April 2018

Candidate Declaration

I hereby certify that the work, which is being presented in the thesis, entitled **Single Nucleotide Polymorphisms in the Base Excision Repair and Double Strand Break Repair pathway genes in Lung Cancer patients and their correlation with clinical outcomes**, in partial fulfillment of the requirements for the award of the degree of **Doctor of Philosophy** in Electronics and Communication Engineering from Thapar University, Patiala is an authentic record of my own work carried out during the period **July 2013 to April 2018** under the supervision of **Dr. Siddharth Sharma, Dr. Navneet Singh** and **Dr. D. Behera**. I have also cited the reference about the text(s)/figure(s)/table(s) from where they have been taken.

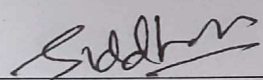
The matter presented in this thesis has not been submitted elsewhere for the award of any other degree or diploma from any institution.

Date: **25-02-2019**

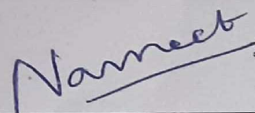


Amrita Singh
Candidate

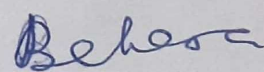
This is to certify that the above statement made by the candidate is correct to the best of our knowledge.



Dr. Siddharth Sharma
Supervisor



Dr. Navneet Singh
Co-Supervisor



Dr. D. Behera
Co-Supervisor

.....*dedicated to my "Father"*

Abstract

Background

DNA repair act as guards to the genomic integrity. Molecular alterations in the repair pathways genes may lead to improper repair and ultimately carcinogenesis.

Objectives

To evaluate the role of single nucleotide polymorphic variants of BER (Base Excision Repair) genes *OGG1* (*OGG1 Ser³²⁶Cys*), *MUTYH* (*MUTYH Gln³²⁴ His (G/C)*), *XRCC1* (*XRCC1 Arg³⁹⁹Gln*, *XRCC1 Arg¹⁹⁴Trp*), *XRCC1 Pro²⁰⁶Pro*, *XRCC1 Arg²⁸⁰His*, *XRCC1 Gln⁶³²Gln*), and DSB (Double Strand Break Repair) genes *XRCC3* (*XRCC3 Thr²⁴¹ Met*), *XRCC4* (*XRCC4 intron 3 I/D*), *XRCC6* (*XRCC6 61G>C*), *XRCC7* (*XRCC7 6721G>T*).

Methodology

Genotyping of genomic DNA was carried out using PCR-RFLP (polymerase chain reaction-restriction fragment length polymorphism) for each polymorphic site under study. Total number of subjects genotyped in case of *OGG1* and *MUTYH* were 656 (330 cases, 326 controls); in *XRCC1* were 655 (330 cases, 325 controls) and in case of *XRCC6* and *XRCC7* were 650 (330 cases, 320 controls) respectively. Following this association analysis was carried out using logistic regression to obtain adjusted odds ratio and significance. Data mining analysis was performed including both Multi-dimensionality reduction (MDR) and Classification and Regression tree (CART) analysis to find the possible interaction between interacting snp-snp and gene-gene. Overall survival analysis was performed using Kaplan meier survival analysis and Cox-regression analysis which gave adjusted hazards ratio.

Results

OGG1 presented a high risk towards lung cancer (*CG*: OR=2.44, **p=0.0003**; *CG+GG*: OR=1.88, **p=0.0093**). On the same lines adenocarcinoma patients in case of *OGG1* were at high risk of acquiring lung cancer (*CG*: OR=4.72, **p=0.0002**; *CG+GG*: OR=3.63, **p=0.0018**). Single allelic carriers for *MUTYH* gene imposed a high risk towards overall lung cancer susceptibility and for all the three histology. Gene-environmental interaction revealed that non-smokers both in case of *OGG1* and *MUTYH* acquired a high risk towards lung cancer. Evaluation of clinicopathological factors revealed that *OGG1* showed a strong correlation with lymph node involved. For the five polymorphic variants of *XRCC1* statistical analysis revealed *XRCC1 Gln⁶³²Gln* (OR=2.67, **p=<0.001**) depicted an overall high risk towards lung cancer. Histological subdivision revealed carriers of mutant genotype in case of *XRCC1 Arg³⁹⁹Gln* imposed a protective effect

towards SQCC subtype. Likewise, mutant genotype in *XRCC1 Pro*²⁰⁶*Pro* implied a protective effect for SCLC subtype (OR=0.29, **p=0.0017**) on the contrary *XRCC1 Gln*⁶³²*Gln* showed a high risk in SQCC diseased group (OR=4.16, **p=<0.0001**). Combination of *XRCC1 Gln*⁶³²*Gln* with other SNPs revealed *XRCC1 Gln*⁶³²*Gln* with *Arg*¹⁹⁴*Trp* (OR=2.10, **p=0.03**) and *Pro*²⁰⁶*Pro* (OR=5.6, **p<0.0004**) increased an overall risk towards lung cancer. Haplotype analysis illustrated haplotype block 11 (CGAGG) carrying minor allele for *XRCC1 206* was associated with the highest risk towards lung cancer on the contrary block 4 (CAGAG) carrying mutant allele for *XRCC1 399* significantly decreased the risk. Multi-dimensionality reduction (MDR) results showed the three factor model comprising *XRCC1 206*, *632*, *280* as the best model (CVC=10, prediction error=0.34). Further Classification and Regression tree (CART) analysis revealed terminal node 1 carrying mutant genotype of *XRCC1 632* and wild type of *XRCC1 280* represented the highest risk group. The DSB gene *XRCC4 intron3* did not show any significant role in modulating the susceptibility towards lung cancer. However, when association analysis was carried out for smokers and non-smokers, the non-smoker individuals with the combined variant genotype were found to be at two times higher risk of developing cancer (OR=1.9, **p=0.04**). The statistical evaluation of *XRCC6 61C>G* and *XRCC7 6721G>T* did not show any significant impact in modulating the susceptibility towards lung cancer. SCLC (Small cell lung cancer) subtype posed an amplified risk towards lung cancer in case of *XRCC7 6721G>T* (OR=4.11, **p=0.0040**). Gene-environment interaction analysis revealed that non-smokers with heterozygous genotype (*CG*) in case of *XRCC6 61C>G* showed a strong protective effect (OR=0.38, **p=0.01**) towards lung cancer. Survival analysis for stratified chemotherapeutic drug regimens revealed, administration of cisplatin/carboplatin+ pemetrexed for *OGG1Ser*³²⁶*Cys* showed a better survival (MST *CG* vs. *CC*: 9.1 vs. 0.56, **p=<0.0001**; HR =0.051, **p=0.0025**). Whereas, *MUTYH Gln*³²⁴*His* showed a smaller survival for mutant genotype (*CC*)(MST *CC* vs. *GG*: 4.0 vs. 9.4, p=0.05; HR= 1.75, p=0.26). The mutant genotype for *XRCC1 399* is observed to have a better survival (MST=9.6). Histological stratification did not reveal any association for any SNP except for SCLC subtype in *XRCC1 632* with an increased death rate (HR=3.08, **p=0.02**). On stratification according to chemotherapy regimen administered; cisplatin/carboplatin+docetaxel was observed to increase survival for *XRCC1 399* mutant genotype (H.R=0.26, **p=0.05**). Cisplatin/carboplatin+irinotecan increased survival in both heterozygotes and combined variants (HR=0.22, **p=0.014**; H.R=0.23, **p=0.012**). Subjects with SCLC subtype for polymorphic variant *XRCC6 61C>G* revealed a poor prognosis as high death rate was obtained. Multiple gene interaction, showed that the three factor model with CVC 10/10 and prediction error =0.33 was the best interacting model determining the lung cancer risk. The second best model was the

four factor model (CVC=9/10, Prediction error=0.29). In CART analysis *XRCC1 632 (H+M)* came out to be the significant factor, *XRCC1 632(H+M)/MUTYH 324(H+M)* *XRCC1 206(W)* showed a twenty fold risk of acquiring lung cancer. Survival tree analysis depicted *OGG1Ser³²⁶Cys* as the major contributor.

Conclusion

This preliminary study provides an insight into the potential of genetic differences in BER and DSBRR genes to modulate lung cancer susceptibility. BER genes *OGG1* and *MUTYH* were certainly modifying risk of lung cancer in North Indian population. *XRCC1 632* acted as an important. risk modifier of lung cancer. Furthermore, polymorphic *XRCC6* and *XRCC7* were not modifying both overall lung cancer susceptibility and survival.

List of Publications

International Journal

1. Amrita Singh, N Singh, D Behera and S Sharma, “*Association and multiple interaction analysis among five XRCC1 polymorphic variants in modulating lung cancer risk in North Indian population*”, DNA Repair, volume 47, pages 30-41, Nov 2016. Impact Factor:3.9
2. Amrita Singh, N Singh, D Behera and S Sharma, “*Polymorphism in XRCC1 gene modulates survival and clinical outcomes of advanced North Indian lung cancer patients treated with platinum-based doublet chemotherapy*”, Medical Oncology, volume 34, issue 4, page 64, Apr 2017. Impact Factor:2.48
3. Amrita Singh, N Singh, D Behera and S Sharma, “*Genetic Investigation of Polymorphic OGG1 and MUTYH Genes Towards Increased Susceptibility in Lung Adenocarcinoma and its Impact on Overall Survival of Lung Cancer Patients Treated with Platinum Based Chemotherapy*”, Pathology & Oncology Research, pages 1-14, Dec 2017. Impact Factor:1.73
4. Amrita Singh, N Singh, D Behera and S Sharma, “*Role of polymorphic XRCC6 (Ku70)/XRCC7 (DNA-PKcs) genes towards susceptibility and prognosis of lung cancer patients undergoing platinum based doublet chemotherapy*”, Molecular biology reports, pages 1-9, Feb 2018. Impact Factor:2.02
5. N Sharma, A Singh, N Singh, D Behera and S Sharma, “*Genetic polymorphisms in GSTM1, GSTT1 and GSTP1 genes and risk of lung cancer in a North Indian population*”, Cancer Epidemiology, volume 39(6), pages 947-955, 2015. Impact Factor:2.644.

Poster presented in Conferences

1. Amrita Singh, N Singh, D Behera and S Sharma “*Role of polymorphic variants of BER and DSB pathway genes in modulating lung cancer susceptibility and prognosis of North Indian Population*”, IASLC 18th World Conference on Lung Cancer, 14th-18th Oct, 2017, Japan.
2. A Singh, N Sharma, N Singh, D Behra, S Sharma “*Genetic polymorphism of Phase II (GSTM1, GSTT1, GSTP1) Detoxification genes with risk for Occurrence Of*

Lung Cancer and its association with Clinico-pathological features In North Indian Population", 4th Annual Convention of Indian Association for Cancer Research, IACR Feb, 2015, Jaipur.

Acknowledgements

This thesis is the end of my journey in obtaining my Ph.D. I have not traveled in a vacuum in this journey. This thesis has been kept on track and been seen through to completion with the support and encouragement of numerous people including my well wishers, my friends, colleagues and the two institutions Thapar University and PGIMER Chandigarh. I would like to thank each and everyone who made this thesis possible and an unforgettable experience for me. It is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study and made it an unforgettable experience for me. At the outset, i am extremely grateful to my main supervisor **Dr. Siddharth Sharma**, Associate Professor for his guidance, and all the useful discussions and brainstorming sessions, especially during the difficult conceptual development stage. His deep insights helped me at various stages of my research. I also remain indebted for him understanding and support me during the times when I was really down and depressed due to personal family problems. His determined and enthusiastic nature has always inspired me to be consistent in my work. Words would be less to express my thanks to him.

My sincere gratitude is reserved for my co-guide **Dr. Navneet Singh**, Additional Prof, Department of Pulmonary Medicine, PGIMER for his invaluable insights and suggestions. His grounded personality has always been an inspiration to look forward to. His attitude to easily solve any problem has always motivated me to stay calm and hard working in my work. I cannot thank him enough in words. I remain amazed that despite his busy schedule, he was able to go through the final draft of my thesis and meet me in less than a week with comments and suggestions on almost every page. He is an inspiration.

I am also obliged to my other co-guide **Dr. Digamber Behera**, Head of Department, Department of Pulmonary Medicine, PGIMER who gave this opportunity to have access to the hospital and encourage research.

I would like to express my gratitude to **Dr. Moushumi Ghosh** Professor and Head of Department, for she ensured that infrastructure is never a problem to the research.

I thank my Doctoral committee members, **Dr. Vikas Handa**, **Dr. Manoj Barnwal**, **Dr. Amjad Ali** for their helpful suggestions and comments during my progress report presentations.

I would also like to thank my lab mates Shweta Lawania, Kinshuk Sharma who have helped me in my work. I would like express my joy working with all M.Sc and M.tech.

students during my research work.

I am also indebted to my friends Shikha Suman, Twinkle Sacchan, Nikhil Khurana, Shipra Singh for being there to listen when i needed an ear and keeping faith in me.

Words cannot express the feelings I have for my parents and my in-laws for their constant unconditional support - both emotionally and financially. I would not be here if it not for them. I dedicate this work of mine to my father Late Shri Satendra Singh Chaudhary. He is no more with us but his presence is always realized. His words "Work hard and live your dreams" always echo in my ears and inspires me.

I expand my thanks to Rajesh Chhiber, Ruchi Chhiber, Himangini Chibber for their constant support and providing a homely environment at a place away from home. I would also like to thank Dr. Painam Balveer, Dr. Ranvir Singh Panwar, Dr. Suresh Kumar, Dr. Shelly Teotia, Prashant Balyan, Prashant Teotia.

Siblings Deeksha Chaudhary and Shubham Chaudhary are always an assets, their immense love and faith were always a driving source to me.

There are times when you are at verge to fall into pieces but the presence of one single inspiration can always keep you going such was my husband **Dr. Pradeep Teotia**. He has been a constant source of strength and inspiration. I can honestly say that it was only his determination and constant encouragement that ultimately made it possible for me to see this thesis through to the end.

At the last I would like to specially mention my son **Viaan Teotia** who is a source of immense energy and freshness every time i see him.

Amrita Singh

Table of Contents

Title	Page No.
Abstract	iii
List of Publications	vi
Table of Contents	x
List of Figures	xvii
List of Tables	xxi
List of Notations	xxvi
List of Abbreviations	xxviii
Chapter 1 Introduction	1
1.1 Introduction	1
Chapter 2 Literature Survey	9
2.1 Lung Anatomy	9
2.2 Lung cancer	10
2.2.1 Histology classification of lung cancer	11
2.2.2 TNM classification	14
2.2.3 Epidemiology of Lung cancer-Worldwide	17
2.2.4 Epidemiology of Lung cancer-India	18
2.2.5 Risk factors	18
2.3 Lung cancer is Age/Gender based	21
2.4 Lung cancer in Never smokers	22
2.4.1 Lung Carcinogenesis	23
2.4.2 Molecular alterations in Lung carcinogenesis	24
2.4.3 Diagnosis of Lung cancer	27
2.4.4 Lung cancer Treatment	28
2.5 DNA damage and Repair	29
2.5.1 DNA damage Repair pathways	31
2.5.2 DNA repair and Carcinogenesis	37

2.5.3	Genetic variabilities in BER genes: <i>OGG1</i> , <i>MUTYH</i> and <i>XRCC1</i> and their correlation with lung cancer predisposition	38
2.5.4	Genetic variabilities in DSB genes <i>XRCC4</i> , <i>XRCC6</i> , <i>XRCC7</i> and their correlation with lung cancer predisposition	41
2.5.5	Polymorphic DNA repair genes (<i>OGG1</i> , <i>MUTYH</i> , <i>XRCC1</i> , <i>XRCC4</i> , <i>XRCC6</i> , <i>XRCC7</i>) on Indian population in various cancer	42
2.5.6	Genetic variants of DNA repair genes in Lung cancer in Indian population	44
2.6	DNA repair and survival	44
2.6.1	Association of genetic variants of DNA repair genes with overall survival and prognosis of Lung cancer patients	45
Chapter 3	Gaps in study and objectives	47
3.1	Study population	47
3.2	Objectives of the study	47
Chapter 4	Material and Methods	49
4.1	Study population	49
4.2	Chemotherapeutic regimens	50
4.3	Follow up and response assessment	50
4.4	Genomic DNA isolation	51
4.5	Qualitative and Quantitative estimation of DNA	52
4.6	Genotyping of polymorphic sites of BER and DSB genes	52
4.6.1	Genotyping of BER genes	52
4.6.2	Restriction Fragment Length Polymorphism	55
4.6.3	Polyacrylamide gel electrophoresis	55
4.6.4	Silver staining	55
4.7	Genotyping of DSB genes	56
4.8	Statistical analysis	58
Chapter 5	Results	61
5.1	Genomic DNA Isolation	61
5.2	PCR amplification	61
5.3	Genotyping of polymorphic variants by RFLP	61
5.3.1	Genotyping for BER genes	61
5.3.2	Genotyping for DSB genes	66
5.4	Characteristics of the study	68
5.4.1	Age based distribution of the study group	68

5.4.2	Distribution of gender of the study group	69
5.4.3	Distribution of smoking status of the study group	69
5.4.4	Distribution of Lung cancer patients on the basis of clinicopathological features of the study group:	71
5.4.5	Distribution of Lung cancer study subjects on the basis of histology	72
5.4.6	Distribution of Lung cancer study subjects on the basis of ECOG and KPS	72
5.4.7	Distribution of study subjects based on Chemotherapeutic regimens administered	73
5.5	Role of polymorphic sites of BER genes <i>OGG1</i> , <i>MUTYH</i> in lung cancer susceptibility	75
5.5.1	Distribution of study subjects with polymorphic <i>OGG1</i> and <i>MUTYH</i> gene	75
5.5.2	Polymorphic <i>OGG1</i> , <i>MUTYH</i> and their role in modifying lung cancer susceptibility-Overall	77
5.5.3	Genotypic and allelic distribution of polymorphic <i>OGG1</i> , <i>MUTYH</i> gene and lung cancer susceptibility based on histology	78
5.5.4	Effect of smoking status on association of polymorphic <i>OGG1</i> , <i>MUTYH</i> and lung cancer risk	82
5.5.5	Relation between clinicopathological characteristics and polymorphic <i>OGG1</i> and <i>MUTYH</i> distribution	84
5.5.6	Combinatorial assessment of polymorphic <i>OGG1</i> and <i>MUTYH</i> towards lung cancer predisposition	88
5.6	Role of polymorphic sites of <i>XRCC1</i> gene in lung cancer susceptibility .	90
5.6.1	Demographic details of lung cancer study subjects and controls with polymorphic <i>XRCC1</i> gene	90
5.6.2	Genotypic and allelic frequencies of <i>XRCC1</i> variants and their association with overall lung cancer risk	92
5.6.3	Genotypic and allelic frequency of <i>XRCC1</i> polymorphic gene and its association with lung cancer risk on the basis of histological subtypes	95
5.6.4	Effect of smoking on polymorphic <i>XRCC1</i> variants and lung cancer risk associated	101
5.6.5	Relationship of genetic variants in <i>XRCC1</i> gene with clinicopathological parameters	105
5.6.6	Combinatorial analysis of risk associated with different polymorphic sites of <i>XRCC1</i> gene	111

5.6.7	Combinatorial analysis of risk associated with triple combinations of different polymorphic sites of <i>XRCC1</i> gene	114
5.6.8	Haplotypes and Linkage disequilibrium analysis of <i>XRCC1</i> variants and their association with lung cancer susceptibility	116
5.6.9	Multifactor Dimensionality Reduction (MDR) analysis for <i>XRCC1</i> variants	118
5.6.10	Classification and Regression Tree (CART) analysis for five polymorphic variants of <i>XRCC1</i> gene	119
5.7	Relationship of polymorphic variants of DSB repair genes <i>XRCC6</i> , <i>XRCC7</i> in lung cancer susceptibility	123
5.7.1	Demography of lung cancer patients and controls with polymorphic <i>XRCC6</i> and <i>XRCC7</i>	123
5.7.2	Genotypic and allelic distribution of polymorphic <i>XRCC6</i> and <i>XRCC7</i> variants and their association with lung cancer risk	125
5.7.3	Genotypic and allelic distribution of <i>XRCC6</i> and <i>XRCC7</i> genetic variants and the stratified association analysis based on histological subtypes	126
5.7.4	Genotypic distribution of <i>XRCC6</i> and <i>XRCC7</i> genetic variants based on smoking status and its association towards risk of developing lung cancer	129
5.7.5	Genetic variations in <i>XRCC6</i> and <i>XRCC7</i> genes and its relationship with clinicopathological parameters	131
5.7.6	Modulation of lung cancer risk due to cumulative effect of <i>XRCC6</i> and <i>XRCC7</i> polymorphism	135
5.8	Relationship of polymorphic variants of DSB repair gene <i>XRCC4</i> in lung cancer susceptibility	137
5.8.1	Demography of lung cancer subjects and controls with polymorphic <i>XRCC4</i>	137
5.8.2	Allelic and genotypic distribution of <i>XRCC4</i> gene and its association towards lung cancer risk	138
5.8.3	Genotypic and allelic distribution of <i>XRCC4</i> variant and risk towards lung cancer based on histology	139
5.8.4	Genotypic distribution of <i>XRCC4</i> intron 3 variant based on smoking status and its association with risk of developing lung cancer .	141
5.9	Survival analysis of BER genes <i>OGG1</i> , <i>MUTYH</i> and <i>XRCC1</i>	143

5.9.1	Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of polymorphic <i>OGG1</i> and <i>MUTYH</i> genes	143
5.9.2	Association of polymorphic <i>OGG1</i> , <i>MUTYH</i> genetic variants with overall survival of lung cancer patients	144
5.9.3	Cox-regression analysis of the factors associated with overall survival of <i>OGG1</i> and <i>Mutyh</i> in lung cancer patients	148
5.9.4	Association of polymorphic <i>OGG1</i> , <i>MUTYH</i> and overall survival according to histological subtypes of lung cancer patients	148
5.9.5	Association of <i>OGG1 Ser³²⁶Cys</i> and <i>MUTYH Gln³²⁴His</i> polymorphism with treatment response towards chemotherapeutic treatment and their correlation with objective response	151
5.9.6	Overall survival of lung cancer patients according to different chemotherapeutic regimens in relevance to polymorphic <i>OGG1</i> and <i>MUTYH</i> genes	152
5.10	Survival analysis of BER gene <i>XRCC1</i>	157
5.10.1	Clinical characteristics of patients with polymorphic <i>XRCC1</i> gene	157
5.10.2	Allelic and genotypic distribution of <i>XRCC1</i> genetic variants and their association with overall survival of lung cancer patients	159
5.10.3	Cox-regression analysis of the factors associated with overall survival of <i>XRCC1</i> polymorphic variants	165
5.10.4	Association of <i>XRCC1</i> polymorphisms and survival according to histological subtypes of lung cancer patients	166
5.10.5	Association of genetic determinants of <i>XRCC1</i> and overall survival of lung cancer patients treated with different chemotherapeutic regimens	172
5.10.6	Genetic determinants of <i>XRCC1</i> and their correlation with clinical stage and performance status in relation to survival	179
5.10.7	Combinatorial effect of double combinations of <i>XRCC1</i> SNPs on overall survival of lung cancer patients	190
5.10.8	Combined effect of triple combinations of <i>XRCC1</i> SNPs on overall survival	194
5.10.9	Relationship of <i>XRCC1</i> polymorphism and their correlation with objective response after chemotherapy	197
5.11	Survival analysis of DSBR genes <i>XRCC3</i>	198
5.11.1	Clinical characteristics of lung cancer subjects with polymorphic <i>XRCC3</i> genotype	198

5.11.2	Polymorphic <i>XRCC3</i> <i>T241M</i> and Survival- Overall and histology based	200
5.11.3	<i>XRCC3</i> polymorphism and overall survival based upon chemotherapy regimens	203
5.11.4	Role of <i>XRCC3</i> polymorphism and its outcome with objective response	204
5.12	Survival analysis of DSB repair gene <i>XRCC4</i>	205
5.12.1	Clinical characteristics of lung cancer patients with polymorphic <i>XRCC4</i>	205
5.12.2	Association of <i>XRCC4</i> variant with overall survival of lung cancer patients	206
5.12.3	Association of <i>XRCC4</i> polymorphism and overall survival according to chemotherapy regimens given to lung cancer patients . . .	209
5.12.4	Association of <i>XRCC4</i> polymorphism and its correlation with objective response	212
5.13	Survival analysis of DSB repair genes <i>XRCC6</i> and <i>XRCC7</i>	212
5.13.1	Clinical characteristics of subjects with polymorphic <i>XRCC6</i> and <i>XRCC7</i>	212
5.13.2	Association of polymorphic <i>XRCC6</i> , <i>XRCC7</i> with overall survival of lung cancer patients	214
5.13.3	Cox-regression analysis of the factors associated with overall survival of <i>XRCC7</i>	217
5.13.4	Association of polymorphic <i>XRCC6</i> , <i>XRCC7</i> and survival according to histological subtypes of lung cancer	217
5.13.5	Role of polymorphic <i>XRCC6</i> and <i>XRCC7</i> gene and overall survival according to chemotherapy regimens administered	221
5.13.6	Association of <i>XRCC6</i> <i>61C>G</i> and <i>XRCC7</i> <i>6721G>T</i> polymorphisms with treatment response for different chemotherapeutic regimens and also their correlation with objective response	226
5.14	Relationship of multiple gene interaction analysis among (<i>OGG1</i> , <i>MUTYH</i> , <i>XRCC1</i> , <i>XRCC4</i> , <i>XRCC6</i> , <i>XRCC7</i>) with lung cancer risk	227
5.14.1	Demographic characteristics	227
5.14.2	Genotypic distribution of SNP-SNP combinations of polymorphisms between BER and DSB repair genes and their association with lung cancer predisposition	228
5.14.3	MDR analysis of BER and DSB repair pathway genes	239
5.14.4	CART analysis of BER and DSB repair genes	243

5.14.5	Association of unfavourable genotypic combinations of BER (<i>OGG1</i> , <i>MUTYH</i> , <i>XRCC1</i>) and DSB (<i>XRCC3</i> , <i>XRCC6</i> , <i>XRCC7</i>) genes with overall survival of lung cancer patients	253
5.14.6	High order gene-gene interactions for prognosis of lung cancer patients using survival tree analysis	261
Chapter 6	Discussion	265
6.1	Discussion	265
6.1.1	BER genes <i>OGG1</i> and <i>MUTYH</i> and its role in lung cancer predisposition	266
6.1.2	Polymorphism in the gene <i>XRCC1</i> and its role towards lung cancer predisposition	269
6.1.3	DSBR genes (<i>XRCC4</i> , <i>XRCC6</i> and <i>XRCC7</i>) polymorphism and lung cancer predisposition	275
6.1.4	DNA repair genes and their role in overall survival and objective response of lung cancer patients receiving platinum based chemotherapy	278
6.1.5	Gene-gene interactions in lung cancer predisposition and survival	285
Chapter 7	Conclusions	290
7.1	Future Prospects of the study	291
References		292
Appendix		326
A.1	Ethics Committee Report	326
A.2	PATIENT INFORMATION SHEET	327
A.3	III	330

List of Figures

Figure No.	Title	Page No.
2.1	Microscopic images of different Histological subtypes of Lung Carcinoma (A) Squamous cell carcinoma (B) Small cell lung cancer (C) Adenocarcinoma (D) Large cell carcinoma (Source:(Weiss, 2005))	13
2.2	Lung cancer trend in four metro cities in 24 years (Source:(<i>LUNG CANCER Focus:INDIA</i> , 2007))	19
2.3	Framework for lung carcinogenesis	24
2.4	Molecular changes during lung cancer pathogenesis (Source:(Panov, 2005))	25
2.5	Specific DNA damage are responded by specific pathways (Source:(Dixon and Kopras, 2004))	31
2.6	Base excision repair pathway	33
2.7	Homologous recombination pathway (Source:(Jackson, 2002))	35
2.8	Non-homologous end joining pathway (Source:(Jackson, 2002))	36
2.9	Flowchart to cellular response to DNA damage	38
5.1	Representative example of agarose gel electrophoresis showing genomic DNA bands. (a) Genomic DNA of lung cancer cases. Lane1 represents the DNA ladder, Lane 2-7 represents the genomic DNA in lung cancer cases (b) Genomic DNA of lung cancer controls. Lane1 represents the DNA ladder, Lane 2-8 represents the genomic DNA in lung cancer controls . .	62
5.2	Representative examples of agarose gel electrophoresis showing PCR-RFLP of BER genes (a) <i>OGG1 Ser³²⁶Cys</i> (b) <i>MUTYH Gln³²⁴His</i> (c) <i>XRCC1 399</i> (d) <i>XRCC1 194</i>	64
5.3	Representative examples of agarose gel electrophoresis showing PCR-RFLP of BER genes (a) <i>XRCC1 206</i> (b) <i>XRCC1 632</i> (c) <i>XRCC1 194</i>	65
5.4	Representative examples of agarose gel electrophoresis showing PCR-RFLP of DSB genes (a) <i>XRCC 3</i> (b) <i>XRCC 4</i> (c) <i>XRCC 6</i> (d) <i>XRCC 7</i> . . .	67
5.5	Age wise distribution of the two genders	68
5.6	Gender based distribution of the study	69
5.7	Smoking wise distribution of the two genders in cases	70
5.8	Smoking wise distribution of the two genders in controls	70
5.9	Stage based distribution of the study subjects	71
5.10	Histological distribution of Lung cancer cases	72

5.11	KPS distribution of Lung cancer cases	73
5.12	ECOG distribution of Lung cancer cases	74
5.13	Distribution of Lung cancer subjects according to drug regimens administered	74
5.14	Pairwise linkage disequilibrium (D') between the five variants of <i>XRCC1</i> gene. White color depicts the minimum LD	117
5.15	The multifactor dimensionality reduction (MDR) interaction dendrogram. <i>X1, XRCC1 399; X2, XRCC1 194; X3, XRCC1 206; X4, XRCC1 632;</i> <i>X5, XRCC1 280</i>	119
5.16	CART analysis in polymorphic <i>XRCC1</i> . W= homozygous wild type geno- type, M= heterozygous + homozygous variant genotype	122
5.17	Kaplan Meier Overall Survival Curve for (a) <i>OGG1 Ser³²⁶Cys</i> (b) <i>MUTYH</i> <i>Gln³²⁴His</i>	146
5.18	Kaplan Meier Overall Survival Curve for <i>MUTYH Gln³²⁴His</i> treated with (a) Cisplatin/Carboplatin + Docetaxel, for <i>OGG1 Ser³²⁶Cys</i> treated with (b) Cisplatin/Carboplatin + Irinotecan	154
5.19	Kaplan Meier Overall Survival Curve for <i>OGG1 Ser³²⁶Cys</i> treated with (a) Cisplatin/Carboplatin + Pemetrexed for <i>OGG1 Ser³²⁶Cys</i> treated with (b) Cisplatin/Carboplatin + Pemetrexed	155
5.20	Kaplan Meier Overall Survival Curve for (a) <i>XRCC1 Arg³⁹⁹Gln</i> (GG:reference), (b) <i>XRCC1 Arg¹⁹⁴Trp</i> (CC:reference), (c) <i>XRCC1 Pro²⁰⁶Pro</i> (AA:reference) for lung cancer patients. H.R:Hazard ration, MST:Median survival time .	161
5.21	Kaplan Meier Overall Survival Curve for (a) <i>XRCC1 Gln⁶³²Gln</i> (GG:reference) (b) <i>XRCC1 Gln²⁸⁰His</i> (GG:reference) for lung cancer patients. H.R:Hazard ration, MST:Median survival time	162
5.22	Kaplan Meier survival Curves for <i>XRCC1</i> variants in lung cancer patients stratified according to histology (a) <i>Arg³⁹⁹Gln</i> (GG:reference) in ADCC pa- tients (b) <i>Pro²⁰⁶Pro</i> (AA:reference) in SCLC patients(c) <i>Gln⁶³²Gln</i> (GG:reference) in SCLC paients. H.R:Hazard ration, MST:Median survival time	169
5.23	Kaplan Meier Overall Survival Curve for <i>XRCC1 Arg³⁹⁹Gln</i> treated with (a) Cisplatin/Carboplatin+Docetaxel (b) Cisplatin/Carboplatin+Irrinotecan (c) Cisplatin/Carboplatin+Pemetrexed	174
5.24	Kaplan Meier Overall Survival Curve for <i>XRCC1 Arg¹⁹⁴Trp</i> treated with (a) Cisplatin/Carboplatin+Docetaxel (b) Cisplatin/Carboplatin+Irrinotecan (c) Cisplatin/Carboplatin+Pemetrexed	175
5.25	Kaplan Meier Survival Curve for <i>XRCC4</i> intron3 I/D variant in lung cancer subjects treated with (a) Cisplatin/Carboplatin +Docetaxel (II:reference) (b)Cisplatin/Carboplatin +Pemtrexed (II:reference)	210

5.26	Kaplan Meier Overall Survival Curve for (a) <i>XRCC6 61C>G</i> (CC:reference)(b) <i>XRCC7 6721G>T</i> (GG:reference). H.R=Hazard rate, MST=Median survival time	215
5.27	Kaplan Meier Overall Survival Curve for XRCC6 and XRCC7 variants in lung cancer subjects stratified according to their histology (a) <i>XRCC6 61C>G</i> (CC:reference) in SCLC subjects (b) <i>XRCC7 6721 G>T</i> (GG:reference) in SCLC subjects	219
5.28	Kaplan Meier overall survival Curve for <i>XRCC6 61C>G</i> subjects treated with (a) Cisplatin / Carboplatin + Docetaxel (b) Cisplatin / Carboplatin + Irinotecan	223
5.29	Kaplan Meier overall survival Curve for <i>XRCC76721 G>T</i> subjects treated with (a) Cisplatin / Carboplatin + Docetaxel (b) Cisplatin / Carboplatin +I rinotecan	224
5.30	Multifactor dimensionality reduction (MDR) interaction dendrogram between BER (<i>OGG1, MUTYH, XRCC1</i>) and DSB genes (<i>XRCC4, XRCC6, XRCC7</i>) X1-XRCC1 399, X2-XRCC1 194, X3-XRCC1 206, X4-XRCC1 632, X5-XRCC1 280, X6-XRCC6, X7-XRCC7, X8-MUTYH, X9-OGG1, X10-XRCC4 (a)In all cases and controls best three factor model (b)In all cases and controls best four factor model	240
5.31	Multifactor dimensionality reduction (MDR) interaction dendrogram between BER (<i>OGG1, MUTYH, XRCC1</i>) and DSB genes (<i>XRCC4, XRCC6, XRCC7</i>):X1-XRCC1 399, X2-XRCC1 194, X3-XRCC1 206, X4-XRCC1 632, X5-XRCC1 280, X6-XRCC6, X7-XRCC7, X8-MUTYH, X9-OGG1, X10-XRCC4 (a) For four factor model in ADCC subjects (b) For three factor model in ADCC subjects	241
5.32	Multifactor dimensionality reduction (MDR) interaction dendrogram between BER (<i>OGG1, MUTYH, XRCC1</i>) and DSB genes (<i>XRCC4, XRCC6, XRCC7</i>): X1-XRCC1 399, X2-XRCC1 194, X3-XRCC1 206, X4-XRCC1 632, X5-XRCC1 280, X6-XRCC6, X7-XRCC7, X8-MUTYH, X9-OGG1, X10-XRCC4 in smokers	242
5.33	CART analysis in polymorphic DNA repair pathway genes. W= homozygous wild type genotype, H+M= heterozygous + homozygous variant genotype	249
5.34	CART analysis in polymorphic DNA repair pathway genes. W= homozygous wild type genotype, H+M= heterozygous + homozygous variant genotype	250

5.35	CART analysis in polymorphic DNA repair pathway genes. W= homozygous wild type genotype, H+M= heterozygous + homozygous variant genotype	263
5.36	Overall survival according to different nodes identified by survival tree analysis	264

List of Tables

Table No.	Title	Page No.
2.1	TNM stage grouping (Source:(Weissferdt, 2012))	15
2.2	Tumor node and metastasis Descriptors (Source:Rusch (2015))	16
4.1	List of Primers and restriction enzymes used in the study	54
4.2	List of Primers and restriction enzymes used in the study	57
5.1	Demographic Characteristics among cases and controls	76
5.2	Assessment of OGG1 and MUTYH risk towards Lung cancer susceptibility	78
5.3	Polymorphic <i>OGG1</i> and <i>MUTYH</i> towards lung cancer development based on Histology	81
5.4	Genotypic distribution of OGG1 and MUTYH based on smoking status and its association with risk of Lung cancer	83
5.5	Relation between clinicopathological factors and <i>OGG1</i> and <i>MUTYH</i> (Stage and primary tumor extension)	86
5.6	Relation between clinicopathological factors and <i>OGG1</i> and <i>MUTYH</i> (Lymph node and metastasis)	87
5.7	Combinatorial assessment of OGG1 and MUTYH towards Lung cancer for overall	89
5.8	Combinatorial assessment of OGG1 and MUTYH towards Lung cancer and Histology for overall	89
5.9	Demographic Characteristics among cases and controls of <i>XRCC1</i>	91
5.10	Polymorphic <i>XRCC1</i> and Lung cancer susceptibility	94
5.11	Genotypic distribution of <i>XRCC1</i> genetic variants and their association with risk of lung cancer based on ADCC histological subtype	98
5.12	Genotypic distribution of <i>XRCC1</i> genetic variants and their association with risk of lung cancer based on SQCC histological subtype	99
5.13	Genotypic distribution of <i>XRCC1</i> genetic variants and their association with risk of lung cancer based on SCLC histological subtype	100
5.14	Genotypic distribution of <i>XRCC1</i> variants based on smoking status and its association with risk of Lung cancer (Smokers)	102
5.15	Genotypic distribution of <i>XRCC1</i> variants based on smoking status and its association with risk of Lung cancer (Non-smokers)	103

5.16	Relation between clinicopathological factors and polymorphic <i>XRCC1</i> (Stage and primary tumor extension)	108
5.17	Relation between clinicopathological factors and polymorphic <i>XRCC1</i> (Lymph node and metastasis)	109
5.18	Genotypic distribution based on different double genotypic combinations of <i>XRCC1</i> variants and their association with lung cancer risk	112
5.19	Genotypic distribution based on triple genotypic combinations of different <i>XRCC1</i> variants and their association with lung cancer risk	115
5.20	Analyzed Haplotype frequencies and risk towards lung cancer due to <i>XRCC1</i> haplotypes	116
5.21	D'/r^2 Values for linkage disequilibrium among five <i>XRCC1</i> SNPs	117
5.22	Multifactor Dimensionality reduction (MDR) analysis of <i>XRCC1</i> variants showing association of high order interactions with Lung cancer	118
5.23	Risk estimates based on CART analysis for <i>XRCC1</i> variants	121
5.24	Demographic characteristics among cases and controls	124
5.25	Genotypic distribution of <i>XRCC6</i> and <i>XRCC7</i> genetic variants and their association with risk of lung cancer	126
5.26	Genotypic distribution of <i>XRCC6</i> and <i>XRCC7</i> genetic variants and their association with risk of lung cancer along with the stratified association based on histology	128
5.27	Genotypic distribution of <i>XRCC6</i> and <i>XRCC7</i> based on smoking status and its association with risk of Lung cancer	130
5.28	Relation between clinicopathological factors and polymorphic <i>XRCC6</i> and <i>XRCC7</i> (Primary Tumor Extension)	133
5.29	Relation between clinicopathological factors and polymorphic <i>XRCC6</i> and <i>XRCC7</i> (Lymph Node and Metastasis)	134
5.30	Genotypic distribution based on different genotypic combinations of <i>XRCC6</i> and <i>XRCC7</i> and their association with lung cancer risk assessment	136
5.31	Combinatorial assessment of <i>XRCC6</i> and <i>XRCC7</i> towards Lung cancer and Histology	136
5.32	Demographic Characteristics among cases and controls of <i>XRCC4</i>	137
5.33	Assessment of <i>XRCC4</i> intron 3 risk towards lung cancer susceptibility	138
5.34	Polymorphic <i>XRCC4</i> intron 3 towards lung cancer based on Histology	140
5.35	Genotypic distribution of <i>XRCC4</i> intron 3 based on smoking status and its association with risk of lung cancer	142
5.36	Patients Characteristics and clinical features	143

5.37	<i>OGG1</i> and <i>MUTYH</i> polymorphic variants and their association with overall survival	147
5.38	Results of cox-regression multivariate analysis for subjects with polymorphic <i>OGG1</i>	148
5.39	Association of <i>OGG1</i> and <i>MUTYH</i> SNPs with survival of lung cancer patients according to Tumor histology ADCC	149
5.40	Association of <i>OGG1</i> and <i>MUTYH</i> SNPs with survival of lung cancer patients according to Tumor histology SQCC	150
5.41	Association of <i>OGG1</i> and <i>MUTYH</i> SNPs with survival of lung cancer patients according to Tumor histology SCLC	150
5.42	Effect of <i>OGG1</i> and <i>MUTYH</i> SNPs on treatment outcome	152
5.43	Association of <i>OGG1</i> and <i>MUTYH</i> SNPs with chemotherapy regimen (Cisplatin/Carboplatin+Docetaxel) and Overall Survival	156
5.44	Association of <i>OGG1</i> and <i>MUTYH</i> SNPs with chemotherapy regimen (Cisplatin/Carboplatin+Irinotecan) and Overall Survival	156
5.45	Association of <i>OGG1</i> and <i>MUTYH</i> SNPs with Chemotherapy Regimen (Cisplatin/Carboplatin+Pemetrexed) and Overall Survival	157
5.46	Patients Characteristics and clinical features of <i>XRCC1</i>	158
5.47	<i>XRCC1</i> polymorphic variants and their association with overall survival	163
5.48	Results of cox-regression multivariate analysis for subjects with polymorphic <i>XRCC1 Arg</i> ³⁹⁹ <i>Gln</i>	165
5.49	Results of cox-regression multivariate analysis for subjects with polymorphic <i>XRCC1 Arg</i> ¹⁹⁴ <i>Trp</i>	165
5.50	Results of cox-regression multivariate analysis for subjects with polymorphic <i>XRCC1 Gln</i> ⁶³² <i>Gln</i>	166
5.51	Association of <i>XRCC1</i> genetic variants with survival of lung cancer patients according to Tumor histology ADCC	170
5.52	Association of <i>XRCC1</i> genetic variants with survival of lung cancer patients according to Tumor histology SQCC	170
5.53	Association of <i>XRCC1</i> genetic variants with survival of lung cancer patients according to Tumor histology SCLC	171
5.54	Association of <i>XRCC1</i> genetic variants with chemotherapy regimen (Cisplatin/Carboplatin+Docetaxel) and survival	177
5.55	Association of <i>XRCC1</i> genetic variants with chemotherapy regimen (Cisplatin/ Carboplatin+Irinotecan) and survival	177
5.56	Association of <i>XRCC1</i> genetic variants with chemotherapy regimen (Cisplatin/ Carboplatin+Pemetrexed) and survival	178

5.57	Association of Polymorphic <i>XRCC1</i> variants with survival of lung cancer patients according to pathological stage IV	181
5.58	Association of Polymorphic <i>XRCC1</i> variants with survival of lung cancer patients according to pathological stage III	182
5.59	Genetic determinants of <i>XRCC1</i> and their association with clinical parameters and outcomes (KPS [100-80])	183
5.60	Genetic determinants of <i>XRCC1</i> and their association with clinical parameters and outcomes (KPS [70-60])	184
5.61	Genetic determinants of <i>XRCC1</i> and their association with clinical parameters and outcomes (KPS [50-40])	185
5.62	Genetic determinants of <i>XRCC1</i> and their association with clinical parameters and outcomes (ECOG [0-1])	188
5.63	Genetic determinants of <i>XRCC1</i> and their association with clinical parameters and outcomes (ECOG [2])	189
5.64	Genetic determinants of <i>XRCC1</i> and their association with clinical parameters and outcomes (ECOG 3-4)	190
5.65	Genotypic combinations and overall survival	192
5.66	Genotypic combinations and overall survival	195
5.67	Effect of <i>XRCC1</i> genetic variants on treatment outcome	197
5.68	Patients Characteristics and clinical features of <i>XRCC3</i>	199
5.69	<i>XRCC 3</i> polymorphic variant and their association with overall survival .	202
5.70	Association of <i>XRCC3</i> genetic variant with survival of lung cancer patients treated with chemotherapy regimen (Cisplatin/Carboplatin+Docetaxel) .	203
5.71	Association of <i>XRCC3</i> genetic variant with survival of lung cancer patients treated with chemotherapy regimens (Cisplatin/Carboplatin+Irinotecan) and Overall Survival	203
5.72	Association of <i>XRCC3</i> genetic variant with survival of lung cancer patients treated with chemotherapy regimens (Cisplatin/Carboplatin+Pemetrexed	204
5.73	Association of <i>XRCC3</i> genetic variant with chemotherapy response . . .	204
5.74	Patients Characteristics and clinical features of <i>XRCC4</i>	205
5.75	<i>XRCC 4</i> polymorphic variant and its association with overall survival . .	208
5.76	Association of <i>XRCC4</i> SNP with chemotherapy regimen (Cisplatin / Carboplatin+Docetaxel) and overall Survival	211
5.77	Association of <i>XRCC4</i> SNP with chemotherapy regimens (Cisplatin / Carboplatin+Irinotecan) overall Survival	211
5.78	Association of <i>XRCC4</i> SNP with chemotherapy regimens (Cisplatin / Carboplatin+Pemterexed) and overall Survival	211

5.79	Effect of <i>XRCC4</i> genetic variants with treatment response	212
5.80	Patients Characteristics and clinical features	213
5.81	<i>XRCC6 61C>G</i> and <i>XRCC7 6721G>T</i> polymorphic variants and their association with overall survival	216
5.82	Results of cox-regression multivariate analysis for subjects with polymor- phic <i>XRCC7 6721G>T</i>	217
5.83	Association of polymorphic <i>XRCC6</i> , <i>XRCC7</i> with Overall Survival ac- cording to Tumor histology ADCC	220
5.84	Association of polymorphic <i>XRCC6</i> , <i>XRCC7</i> with Overall Survival ac- cording to Tumor histology SQCC	220
5.85	Association of polymorphic <i>XRCC6</i> , <i>XRCC7</i> with Overall Survival ac- cording to Tumor histology SCLC	221
5.86	Polymorphic <i>XRCC6</i> and <i>XRCC7</i> and survival for subjects administered Cisplatin/Carboplatin+Docetaxel	225
5.87	Polymorphic <i>XRCC6</i> and <i>XRCC7</i> and survival for subjects administered Cisplatin/Carboplatin+Irinotecan	225
5.89	Effect of Polymorphic <i>XRCC6</i> and <i>XRCC7</i> on treatment response	227
5.90	Demographic Characteristics among cases and controls	228
5.91	Genotypic distribution based on different genotypic combinations and their association with lung cancer risk	233
5.92	Multifactor Dimensionality reduction (MDR) analysis showing association of high order interactions with lung cancer	242
5.93	Overall Risk estimates based on CART analysis (BER and DSBR genes)	245
5.94	Overall risk estimates based on CART analysis (BER and NER genes) among smokers	251
5.95	Overall risk estimates based on CART analysis (BER and NER genes) among non-smokers	252
5.96	Genotypic combinations and overall survival	255
5.97	High order gene-gene interactions using survival tree analysis	262

List of Notations

A	Adenine
α	Alpha
R	Arginine
N	Asparagine
D	Aspartic acid
B	Aspartic acid or Asparagine
bp	Base pair
β	Beta
cm	Centimetre
χ^2	Chi-square
D'	Coefficient of linkage disequilibrium
C	Cytosine
$^{\circ}$	Degree
$^{\circ}C$	Degree Celsius
df	Degree of Freedom
$\geq$	Equals to or greater than
r^2	Extent of linkage disequilibrium
γ	Gamma
E	Glutamic Acid
Q	Glutamine
Z	Glutamine or glutamic acid
G	Gaunine
g	Gram
H	Histidine
h	Hour
HCl	Hydrochloride
I	Isoleucine
$<$	Less than
L	Leucine
l	Litre
K	Lysine
M	Methionine
μg	Microgram
μl	Microliter

μm	Micrometer
μM	Micromolar
mg	Milligram
mg/m^2	Milligram per square meter
mm	Millimetre
min	Minute
M	Molar
$ng/\mu l$	Nanograms per microlitre
%	Percentage
F	Phenylalanine
$\pm$	Plus Minus
P	Proline
s	Second
S	Serine
$AgNO_3$	Silver nitrate
$NaCl$	Sodium Chloride
T	Thymidine
Y	Tyrosine
W	Tryptophan
V	Valine
v	Volume

List of Abbreviations

ADCC	Adenocarcinoma
AMP	Adenosine monophosphate
Arg	Arginine
Asp	Asparagine
β-TrCp	Beta transducin containing protein
BSA	Bovine Serum Albumin
BER	Base Excision Repair
CART	Classification And Regression Tree Analysis
CR	Complete Response
CSC's	Cancer stem cells
CVC	Cross validation consistency
Cys	Cysteine
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribose Nucleoside Triphosphates
DSBR	Double Strand Break Repair
EDTA	Ethylene diamine tetra acetic acid
EMA	Epithelial Membrane Antigen
EGFR	Epidermal growth factor receptor
FDR	False discovery rate
Gln	Glutamine
H	Heterozygote
His	Histidine
HR'	Adjusted hazards ratio
HR	Hazard ratio
HWE	Hardy-Weinberg equilibrium
IASLC	International Association for the Study of Lung Cancer
ICMR	Indian Council Of Medical Research
Ile	Isoleucine
LD	Linkage disequilibrium
M	Mutant
MAF	Minor Allele Frequency
MDR	Multifactor Dimensionality Reduction
m-RNA	Messenger Ribonucleic Acid
MST	Median survival time

NSCLC	Non-Small Cell Lung Cancer
OS	Overall Survival
OR	Odds ratio
OGG1	8-Oxoguanine glycosylase
PAGE	Polyacrylamide Gel Electrophoresis
PCP	Planar cell polarity
PCR	Polymerase Chain Reaction
PGIMER	Postgraduate Institute of Medical Education and Research
Pro	Proline
RBC	Red Blood Cells
RFLP	Restriction Fragment Length Polymorphism
SCF	Stem cell factor
SCLC	Small Cell Lung Cancer
SER	Serine
SNP	Single Nucleotide Polymorphism
SQCC	Squamous Cell Carcinoma
TEMED	N, N, N', N' Tetramethylethylene-1, 2-Diamine
Thr	Threonine
Trp	Tryptophan
TNM	Tumour Node And Metastasis
AOR	Adjusted Odds ratio
VEGF	Vascular Endothelial Growth Factor
W	Wild
WBC	White Blood Cells
WHO	World Health Organization

Chapter 1

Introduction

1.1 Introduction

The term lung cancer refers to bronchogenic carcinoma that has its origin in respiratory epithelium. Lung cancer is increasingly emerging as a fast developing fatal cause for death across the globe. Worldwide, according to a report by WHO lung cancer accounts to 1.38 million cancer deaths (*WHO*, 2018). Continuous increase in lung cancer research in India has focussed light on constant increase on significant lung cancer burden in India. GLOBACON 2012, in its statistical surveys has reported lung cancer ranked fourth in case of incidence rate and has reported nearly 70,275 lung cancer cases. Lung cancer incidence is characteristically ranked second in males compared to sixth in females. Lung cancer is majorly classified as non-small cell lung cancer (85%) with 10 % as small cell lung cancer and remaining 5% with other histology. Lung cancer is not limited to single causative factor. The immense literature and surveys round the world validate that lung cancer is a multifactorial disease where genetic predisposition play an important role. Besides universal factors such as smoking, lifestyle etc. there can be certain factors which are limited to ethnicities. Therefore, to understand the progression of the disease and prognosis its important that these factors are dealt with in detail.

Lung cancer is a consequence of genetic predisposition in constricted association with carcinogenic exposure (Kanwal *et al.*, 2017; Yang *et al.*, 2013; Sellers *et al.*, 1992). At cellular level to counteract the damaging effects of pernicious agents and maintain genome integrity, alteration in the activity of tumor suppressor and proto-oncogenes occur (Sekido *et al.*, 2003). Cell machinery gets activated and cell cycle check points either arrest the cell progression or rectifies the damage. If damage is repaired correctly by DNA repair pathways, cell cycle progresses otherwise it leads to genomic instability giving way to oncogenesis. In this sequence DNA repair system is considered as a safeguard for human genome from constant impertinence implicated by carcinogenic and mutagenic agents such as cigarette/tobacco smoke, environmental exposure, oxidizing agents, ROS species, alkylating agents, ionizing radiations, occupational hazards, etc.

Endogenous factor reactive oxygen species (ROS) produced inside the cell arises due to oxidation, methylation and deamination and constitute important endogenous risk factors. On the other hand factors like smoking, ultraviolet rays, chemical mutagens, air pollution are important exogenous risk factors responsible for the generation of ROS. ROS start acting as a risk factor when an imbalance occurs between biochemical antioxidants and ROS. Oxidative stress has been observed in cancerous cells compared to normal cells therefore it may be related to oncogenic stimulus (Przybylowska *et al.*, 2013). ROS developed in the process are responsible for oxidative damage and development of cytotoxic DNA lesions such as 8-hydroxoguanine (8-oxo-G) and 7, 8-dihydro-8oxo-2-deoxyguanosine. These lesions are studied to be cytotoxic, mutagenic and contribute towards neuro-degeneration, ageing and cancer. The lesion 8-oxo-G can base pair with adenine as well as cytosine thereby responsible for G:C-T:A transversions.

The exogenous factor tobacco smoke has high content of aromatic amines e.g. 2 - Naphthylamine, 4-aminobiphenyl, N-nitrosamines, similarly heavy fuel combustion produces elevated levels of PAH (Polycyclic aromatic hydrocarbon) e.g. B[a]P in the environment (Shihadeh *et al.*, 2015; Chang *et al.*, 2017; Gelboin, 1980). These carcinogens when inhaled are bio-activated by xenobiotic metabolism of cells. During this process, highly lethal electrophilic and free radicals are generated (MacLeod and Tang, 1985). PAH form DNA adducts which are hugely found in respiratory epithelium and act as a precarious factor for induction of lung carcinogenesis (Pohjola *et al.*, 2003; Wiencke *et al.*, 1995). Evolutionary conserved DNA repair pathways have evolved over the time in cells which actively remove the DNA lesions created by mutagens (Houtgraaf *et al.*, 2006). Therefore, lung cancer susceptibility is modulated by the damage removal rate which is genetically determined by DNA repair capacity (DRC). The complex repair signalling pathways act on specific type of damages caused.

Base excision repair pathway (BER) is one such primary sequential pathway that acts on small lesions generated such as base damage caused by ionizing radiations, strong alkylating agents as well as endogenous DNA damaging agents such as ROS species (Krokan and Bjørås, 2013; Sancar, 1995). The damaged base is then removed by a group of enzymes called DNA glycosylases. Glycosylases hydrolyzes the N-glycosyl bond between base and sugar moiety thereby releasing free damaged base and develop abasic site. MUTYH gene (mutY homolog) is another glycosylases which is responsible for the removal of adenine paired with 8-oxo-G or 1,2-dihydro-2-oxo-adenine(2-OH-A) created by ROS species (Sliwinski *et al.*, 2011). Once the damaged base is removed, AP endonucleases like APE1 enzymes are brought to the damaged site and incise the phosphodiester bond, this incision generates 5dRP and 3OH for further synthesis and ligation to occur (Dexheimer, 2013).

The BER pathway can be divided into long and short patch repair pathways depending on if single base is repaired or patch of bases is repaired. *OGG1*, *MUTYH*, *XRCC1* are some important genes of BER pathway.

hOGG1 gene is mapped on chromosome 3 at 3p26.2, and encodes a 345 amino acid product. *hOGG1* gene is highly polymorphic and it has been reported that it has 231 single nucleotide polymorphisms (SNPs) and out of these, twenty five SNPs have reported an amino acid change (Duan *et al.*, 2012) where out of them *Ser*³²⁶*Cys* has been immensely explored and it is located on exon 7 of the gene. The *OGG1* gene is mainly responsible for the removal of mutagenic lesion 8-OxoGua. In never-smokers a high 8-oxodG excreta is observed in urinary excretion with a high risk of lung cancer. The missense polymorphism of *OGG1* i.e. *Ser*³²⁶*Cys* exerted an increased susceptibility to ROS species in an in vitro complementation assays (Loft *et al.*, 2012). Further, human lymphocytes DNA having variant allele of *OGG1* had an elevated level of oxidized guanine compared to DNA with wild or heterozygous genotype (Manini *et al.*, 2009). In another study it was observed non-smokers along with populace within Asian subcontinent carrying homozygous variant allele for *OGG1* were observed to be at increased risk of lung cancer (Vineis *et al.*, 2009). On the same lines *MUTYH Gln*³²⁴*His* is extensively analyzed for its effects on cancer. The gene *MUTYH* is located on chromosome 1 at 1p34.1. It spans a region of 11.2kb. Initial studies revealed that recessive familial form of colorectal cancer (CRC) known as *MUTYH* associated polyposis (MAP) was linked to *MUTYH* gene of base excision repair pathway. It has been also speculated that biallelic germline mutations in *MUTYH* gene were associated with increased colorectal risk (Poulsen and Bisgaard, 2008). In an experiment on mice with null *MUTYH* it was observed that risk was amplified towards carcinogenesis of intestine, liver and lung (Nakabeppu *et al.*, 2004).

The *XRCC1* protein is an important component of BER, because it serves as a scaffold for two other proteins, DNA ligase III and pol and also serves as a single strand break sensor by its interaction with PARP (Ladiges, 2006). *XRCC1* gene has been mapped to 19q 13.2-13.3 position and has 17 exons that spans a genomic distance of 32 kb. At times sequence variants may develop in the numerous genetic participants of the repair pathway due to continuous exposure to carcinogenic agents. Thus it may be plausible that *XRCC1* gene may start behaving as a candidate cancer susceptible gene. Therefore, the polymorphic behavior of the genetic participants may alter the phenotypic DNA repair capacity of individuals. The risk becomes more rigorous due to inheritance of genetic variants at one or more loci. In this respect, *XRCC1* has 30 gene variants that have been identified on exons and promoter region of gene. Out of these variants some are found in 5-UTRs region, some are insertion /deletions, four are synonymous and eight are non-

synonymous SNPs. Out of eight non-synonymous variants three are located at conserved region of *XRCC1*: codon 399(Arg-Gln), codon 194(Arg-Trp), and codon 280(Arg-His) (Shen *et al.*, 1998). The single nucleotide polymorphism (SNP) at codon 399 is found at C terminal domain of BRCT1 (PARP interacting domain). Individuals with variant allele of *XRCC1 Arg³⁹⁹ Gln* have been shown to have higher level of DNA adducts (Lunn *et al.*, 1999) and are sensitive to ionizing radiations (Metsola *et al.*, 2005). *XRCC1 Arg¹⁹⁴ Trp* is located in between N-terminal domain and NLS (nuclear localization domain) whereas *XRCC1 Arg²⁸⁰ His* is located in PCNA binding region (Fan *et al.*, 2004). Variant carriers of *XRCC1 Arg²⁸⁰ His* have impaired the repair capacity due to improper localization of damaged site (Li *et al.*, 2008). The synonymous variant *XRCC1 Pro²⁰⁶ Pro* is located in exon 7 of *XRCC1* gene and interacts with *hOGG1* and *APE1*, whereas *XRCC1 Gln⁶³² Gln* is found in exon 17 and ligase III is the interacting partner.

Damages like double strand breaks (DSBs) are considered the most fatal damages manifested. DSBs are lethal breaks, which if are not repaired correctly or are not repaired would lead to chromosomal aberrations or deletions finally resulting in carcinogenesis along with other deficiency syndromes (Torgovnick and Schumacher, 2015; Kanaar *et al.*, 1998). DSBs are repaired either by Homologous recombination (HR) or Non-homologous end joining (NHEJ). HR requires undamaged sister chromatid as template however NHEJ is involved in the direct ligation of broken ends and therefore is prone to errors. NHEJ is predominantly found in all phases of growth in mammalian cells however HR is limited to late S and G-2 phase. The NHEJ pathway makes use of number of proteins in which the initial step involves the recruitment of Ku70 (*XRCC6*)/ Ku80 (*XRCC5*) heterodimer to the exposed DNA termini. Ku adopts ring shaped structure and acts as a DNA targeting subunit of the DNA-dependent protein kinase catalytic subunit (DNA PKcs or *XRCC7*) which along with Ku forms DNA-PK holoenzyme. DNA-PKcs induce an inward translocation of Ku protein allowing DNA-PKcs to contact opposing DSB ends. Synapsis of DNA-PKcs result in auto-phosphorylation resulting in contact of two DNA termini and finally ligation occur with the help of *XRCC4* and Ligase (Karran, 2000).

Double strand break repair pathway gene *XRCC6* also called as KU70 or G22P1 is located on chromosome 22q13 and encodes the KU70 protein. *XRCC7* also known as PRKDC/HYRC/HYRC1 which is found on chromosome 8q11 (Wang *et al.*, 2008). Genetically modified DNA repair genes are associated with increased genomic instability and ultimately altered DNA repair capacity (DRC). It has been reported that subjects with reduced DRC are more susceptible towards lung cancer risk as studied by phenotypic assays (Spitz *et al.*, 2003). Similarly, lung cancer patients have been reported to have markers of chromosome instability such as high fractional allele loss (FAL) caused by

functional inactivation of DSB repair genes (Tseng *et al.*, 2009). One of the genetic variants of XRCC6/KU70 gene is the 61C>G (rs2267437) that is located in the promoter region. The recognition elements (i.e. Sp1 and AP-2alpha) found in KU70-61G are more in number than the DNA sequence containing KU70-61C (i.e. Sp1), thus implying that this polymorphism has the potential to alter KU70 transcription. This polymorphism in the XRCC6/KU70 gene affects promoter activity. The genetic variants of the XRCC7 i.e. 6721G>T is located in the intron 8 and plays an important role in splicing regulation that causes mRNA instability. Therefore, both XRCC6 and XRCC7 play an important role in the suppression of chromosomal rearrangements and maintenance of genome integrity along with in recognition and repair of double strand breaks.

Germline variations of these two repair pathway genes have been reported in various cancers. There are many reports which have evaluated the role of genetic differences in DNA repair pathways genes towards predisposition of the disease namely OGG1 Ser³²⁶Cys and MUTYH Gln³²⁴His (Kasahara *et al.*, 2008; Sliwinski *et al.*, 2011; Duan *et al.*, 2012), non-synonymous mutation of XRCC1 (Hu *et al.*, 2011; Kim *et al.*, 2002; Chacko *et al.*, 2005; Wang *et al.*, 2015; Stern *et al.*, 2001; Lee *et al.*, 2001; Han *et al.*, 2014; Kang *et al.*, 2015). Apart from these non-synonymous mutations certain synonymous mutation evaluated are XRCC1Pro²⁰⁶Pro and Gln⁶³²Gln they have been analyzed before by [(Mahjabeen *et al.*, 2013; Yin and Wang, 2012; Han *et al.*, 2014)]. The KU70-61C>G and XRCC76721G>T polymorphisms have been studied for their role in various cancers like breast (He *et al.*, 2012; Nasiri *et al.*, 2012), glioma (Wang *et al.*, 2004), renal (Hirata *et al.*, 2006), hepatocellular (Hsieh *et al.*, 2015), digestive (Yang *et al.*, 2011), bladder (Wang *et al.*, 2008).

The alarming fact with regard to lung cancer is that in spite, of intensive effort to improve the modalities of lung cancer diagnosis; this disease has a dismal overall five year survival rate and has accounted to 15% over past many years and this number is much less in India (Globacon, 2012). Most patients at the time of diagnosis are usually at an advanced stage (stage III or IV) or show metastatic disease; thus the window of opportunity for surgical resection is lost thus making curative surgery impossible (Molina *et al.*, 2008). The main conventional treatment for lung cancer is platinum based combination chemotherapy which is one of the most widely and extensively prescribed treatment for advanced lung cancer patients. Combination chemotherapy involves co-administration of platinum based drug along with another non-platinum based drug (*National Comprehensive Cancer Network*, 2016). These drugs form platinum DNA adducts which may cause inter-strand and intra-strand cross-linking that results in distortion of DNA and ultimately inhibit DNA replication (Rosell *et al.*, 2002). Such combination based ther-

apy no matter have increased the overall survival (OS) of patients, however in clinical practice the chemotherapeutic efficacy varies among individuals ranging from 26-60% (Bahl and Falk, 2001). Some patients respond to chemotherapy whereas others develop drug resistance. It is a known fact that genetic factors do play an important role in the development of disease, response for drugs, and its effectiveness thus directly affecting patients survival. Thus all these factors play a role that affects a patients prognosis. It has been speculated that single nucleotide polymorphisms (SNPs) or mutations that occur in DNA repair pathway genes might alter gene expression and DNA repair capacity, therefore it might influence the effectiveness of chemotherapy, prognosis and survival of patients (Abdel-Rahman and El-Zein, 2000).

DNA repair capacity (DRC) has potential to be used as an effective prognostic marker. A number of molecular markers have been reported with prognostic significance but none of them have been used in clinical trials (Sato *et al.*, 2007). Studies have revealed relation between inefficient DNA repair capacity and lung cancer risk (Wei *et al.*, 1996) implying a hypothesis that suboptimal DNA repair capacity might lead to decreased removal of platinum-based adducts and ultimately a better response towards drugs (P Basourakos *et al.*, 2017; Cimino *et al.*, 2013; Shellard *et al.*, 1993). Therefore compromised DRC may behave as a better prognostic factor for improved survival.

The repair pathway genes discussed above for susceptibility are looked for their influence on survival. *XRCC1* protein of BER pathway is critical in repairing damages implicated by damaging agents such as ROS species, alkylating agents, platinum based drugs such as cisplatin and carboplatin (Chen, Zhao and Shi, 2012). Studies have shown that *XRCC1* could bind to platinum containing DNA duplexes (Zhu and Lippard, 2009). NSCLC cell lines when treated with cisplatin, have shown a higher level of *XRCC1* transcript indicating a correlation with cisplatin resistance (Weaver *et al.*, 2005). Thus it has been proposed that the *XRCC1* protein is critical for repairing other types of DNA damage induced by the platinum-based anticancer drugs (such as cisplatin and carboplatin). Therefore it might be plausible that genetic variants in the *XRCC1* gene may modulate the DRC (Au *et al.*, 2004) and hence influence the anticancer effect of platinum based drugs; thus indirectly playing a role in the prognosis and survival of lung cancer patients.

The three common polymorphism of *XRCC1* i.e. *XRCC1 Arg³⁹⁹Gln* (exon10), *XRCC1 Arg¹⁹⁴Trp* (exon6), *XRCC1 Arg²⁸⁰His* (exon 9) (Shen *et al.*, 1998) cause amino acid change which might modify the DNA repair capacity and thereby the response towards the medication. *XRCC1 399Gln* allele is reported to be associated with higher levels of aflatoxin B1 (AFB1-DNA) suggesting the variant allele might impair the repair process (Lunn *et al.*, 1999). *XRCC1 194* was observed to be a significant prognostic factor for

progression free survival (Yoon *et al.*, 2005).

Another important gene lung cancer survival is *XRCC3 Thr²⁴¹Met* which has been studied for its independent role in chemotherapeutic treatment (De las Penas *et al.*, 2006). DNA repair gene *XRCC3* makes an important part of Homologous recombination sub-pathway of Double strand break repair pathway. *XRCC3* is a member of RAD-51-related proteins that take part in homologous recombination. *XRCC3* initiates HR and functions in stabilizing hetero-duplex formed during homologous recombination. Radiated cells showed that *XRCC3* deficient cells were incapable for homologous recombination (Brenneman *et al.*, 2000). The gene *XRCC3* is localized on chromosome 14 and the polymorphism *XRCC3 C¹⁸⁰⁶⁷T* is present on exon 7 of the gene. A transition from C-T occurs resulting in Thr to Met at codon 241 of the gene (R.B. Ji and Hu, 2015). The polymorphism *XRCC3 Thr²⁴¹M* has been studied to be associated with bulky DNA adducts in WBCs (Matullo *et al.*, 2001). *XRCC3 Thr²⁴¹M* is a non-conservative change due to which there may be certain biological implications for activity of enzyme or an interaction with other proteins involved in DNA repair (Gulnaz *et al.*, 2013). The present work utilizes candidate susceptibility gene based approach towards evaluating lung cancer for the polymorphic variants. However, it is evident from the reserve of literature present that genetic factors act synergistically with environmental factors in modulation of cancer. Though genetic constituent acts as an important role player in the process of carcinogenesis (Mucci *et al.*, 2001). The difference in DRC cannot be predicted based on single gene approach for a complex pathways of repair. Even, lung cancer is not governed by alteration of single gene, the immense research elucidate that various signalling pathways amalgamate along with environmental factors in the etiology of cancer. Hence multiple pathway based approach, by considering the cumulative effect of low penetrant genes along with environmental parameters would help in unveiling a clear picture (Neumann *et al.*, 2005). The DNA repair pathways are key role players in maintaining genome integrity, they are complex signalling cascades. Therefore, the interactive potential of multiple enzymes in the repair phenotype is needed to be addressed.

The primary reason for evaluating the sequence variants of DNA repair pathways genes (*OGG1*, *MUTYH*, *XRCC1*, *XRCC3*, *XRCC4*, *XRCC6*, *XRCC7*) and their cumulative effect is that there are very few reports which have evaluated such large number of repair genes together in lung cancer. The analysis states that these genetic variants which form part of genomic guards have important role in lung cancer predisposition and prognosis. In addition, this is the first report evaluating the two synonymous polymorphisms of *XRCC1* in the study *i.e.* *XRCC1 Pro²⁰⁶Pro* and *XRCC1 Gln⁶³²Gln*. As far as our knowledge is concerned this is the first report in North Indian population evaluating all

repair genes together. We have made an attempt in the present work to find the high risk posing subgroups within the sample population by stratifying the subjects on the basis of their histology, clinicopathological features, and smoking habits. Haplotypic analysis was also performed for SNPs present on the same gene *XRCC1*. Further, the present study was also designed to explore the high order gene-gene and gene environment interactions. Multifactor dimensionality reduction (MDR) and Classification and regression tree analysis (CART) were used as major tools to find the cumulative effect of all the polymorphic variants and their interaction with environmental factors. The SNPs are also analyzed in the present study to analyze their association with survival of lung cancer subjects in North Indian population.

Chapter 2

Literature Survey

2.1 Lung Anatomy

Respiratory system is divided into two parts airways anatomy and lung anatomy. Airways anatomy include extra-thoracic and intra-thoracic airways. Lung anatomy includes lung parenchyma which is involved in gas exchange at the alveolar level. Lungs are soft spongy pyramid shaped paired organs found in the upper thoracic cavity. Lungs are surrounded by chest wall which is made up of ribs and muscles between ribs. Two lungs are separated by mediastinum where heart and other organs are present (Nagaishi and Nagasawa, 1972). Diaphragm, a thin muscle separates abdomen from respiratory cavity. The lungs extend laterally from the heart to the ribs on both sides of the chest and continue posteriorly towards spine. Lung is cone shaped with inferior end of lungs resting on diaphragm. The base of lungs is concave to follow the contour of the diaphragm. The left lung is smaller than the right lung because 2/3 rd of the heart is located on the left side of the body. Lungs are lobular where left lung has two lobes separated by left oblique fissure. Right lung has three lobes horizontal fissure separates the first two lobes while the middle and lower one are separated by right oblique fissure (OpenStax, 2014).

Lungs are covered by pleural cavity which is doubled layered and constitute visceral and parietal pleura. Lungs are covered with visceral pleura on inner side which is simple squamous epithelium layer. This visceral pleura on upper surface changes to parietal pleura. Between the two layers there is a cavity called pleural cavity which allows for the expanding of lungs during inhalation. There is serous liquid present in between the two layers of pleural cavity which lubricates and help to prevent irritation during breathing (ElaineN.Marieb, 2012). Blood vessels, primary bronchus and nerves enter lungs through a point called hilum. Blood vessels bronchus and nerves together are called "root of the lung".

Air enters the body through nose or mouth and passes to pharynx, larynx and trachea sequentially. Pharynx can be broadly classified into three parts "Nasopharynx", "Oropharynx" and "Laryngopharynx". Nasopharynx is located next to nasal cavity. Its

lined with ciliated pseudostratified columnar epithelial tissue. Oropharynx comes next to Nasopharynx and it provides for passage of food, air, drink (TerryDesJardins, 2008). Its lined with stratified squamous epithelial tissue. Laryngopharynx is the last part of pharynx that connects to larynx through an opening called glottis. It is also lined with stratified squamous epithelial tissue for dealing with swallowing of food, drink and air. Larynx is called the voice box. Its made up of nine cartilages, out of them eight are of hyaline cartilage connective tissue and the ninth cartilage i.e. epiglottis is of elastic cartilage connective tissue. Epiglottis provide for swallowing of food by covering larynx to prevent food from entering larynx. Next air moves to Trachea, the windpipe. Trachea is a membranous structure, made up of 18-20 C-shaped hyaline cartilage connective tissue (Tu.J:Inthavong, 2013). Trachea is held upright with the help of these cartilages, for easy passage of air through it. Smooth muscles are present in the posterior wall of the trachea which assist in contracting and relaxing of trachea. Just before reaching the lungs, trachea splits into right and left bronchi in the mediastinum. Trachea are hollow tubes, made up of hyaline cartilage like trachea and are lined with ciliated pseudostratified epithelium. The hyaline cartilage "C" shaped rings are found in the interior of bronchi and prevent it from collapsing and blocking airflow to the lungs.

Primary bronchi continues to split into secondary and tertiary bronchi respectively till it divides individually for each lobe. This division helps to evenly divide the air flow in each lobe of the lung. Tertiary bronchi are further split into bronchioles and finally bronchioles into terminal bronchioles into respiratory bronchioles and terminating at alveolar duct (Empey, 1978). Alveolar ducts leads to cluster of alveoli called alveolar sacs. Gaseous exchange occur at thin surface of alveolar duct, alveoli and respiratory epithelium which is made of simple squamous epithelium (Dalavanga, 2017).

2.2 Lung cancer

Lung cancer is a state of lungs where cells undergo progressive pathological changes such as pre-neoplastic or premalignant lesions formation. In a normal body without cancer cells, growth is in control by a system of checks and balance on cells, so that cells divide to produce new cells only when needed. Cancer begins with an error or mutation in cell's DNA (Loeb and Loeb, 2000). When these mutations affect cells in a way that there occurs a disruption in the control over cell death and cell growth, a tumor mass may grow. This tumor mass can be "malignant" or "benign" (*Type of Tumors*, 2016). Benign tumors do not spread to other parts of the body and can be removed from body. Tumors become malignant when tumors are able to invade normal tissues. Tumors can

even enter blood or lymphatic system and therefore can reach to other parts of the body (*Lung Cancer*, 2017). This process of spread of tumor to other sites is called metastasis. The most common metastatic sites of lung tumor are adrenal glands, bones, liver and brain (Tamura *et al.*, 2015). Lung cancer primarily arise from epithelial cells, the cells lining the larger and smaller air pathways therefore it can be called as bronchogenic carcinoma (Press, 1991). When cancer arise from pleura its called mesotheliomas.

2.2.1 Histology classification of lung cancer

Lung cancer is classified into two major clinico-pathological groups Small cell lung carcinoma (SCLC) (20%) and Non-small cell lung carcinoma (NSCLC) (80-85%). NSCLCs are further divided into three subtypes. Mesothelioma and carcinoid tumors are also some less known subtype of lung tumor. Smoking is responsible for all types of histological subtype however, SQCC is strongly associated with smoking followed by SCLC. ADCC are majorly observed in non-smokers and women. The difference histological subtypes is influenced by evolution of smoking related-lung cancer over time and geographical locations (Travis and Harris, 2004). There occurs a heterogeneity in the subtypes of lung cancer, comprising cells of different histology *Lung Cancer* (2019). When a clear major type of cell is present the subtype is named after it and when different types are present in same number the subtype is designated as adeno-squamous carcinoma.

Squamous Cell Carcinoma

Squamous cell carcinomas (SQCC) are observed in smokers in (90%) of cases. SQCC are malignant epithelial cells with a characteristic expression of high molecular weight keratin, cytokeratins and carcino-embryonic antigen (CEA) (Chu and Weiss, 2002). Majority of squamous cell carcinoma occur in close central part of lung (lobar), mainstem or segmental bronchi. Squamous cell carcinoma begin from a precancerous lesion squamous metaplasia or dysplasia in bronchial epithelium. Squamous tumor shows irregular often friable, grey white cut surface with a large area of central necrosis with or without cavity formation as indicated in figure 5.10 (A). SQCC tends to be locally aggressive however metastasis to distant organs is less frequent. There are higher chances of loco-regional recurrence after surgical resection of tumor compared to other cells type. Histopathologically squamous cell carcinoma range from well-differentiated squamous cell neoplasm with keratin formation and intracellular bridges to poorly differentiated with minimal residual squamous cell features. The SQCCs are observed to spread locally within chest directly making use of mediastinal structure. However, carcinomas that are poorly differentiated metastasize early and to distant sites (Funai *et al.*, 2003). Histochemically they appear

TF1 negative and CK 5, 6 P63, P40 positive. SQCCs are either diploid or hyperdiploid-aneuploid with mean chromosome number in triploid range. Cytogenetics shows there is characteristic amplification of telomeric 3q regions along with a gain of 3q24 in majority of squamous cell carcinoma along with deletion on short arm of chromosome 3. 84% of SQCC are observed to be EGFR positive whereas HER2 and K-ras expression is rare in SQCC subtype.

Small Cell Carcinoma

Small cell carcinoma arise from epithelial cells that become malignant. Most of the cases of SCLC show neuroendocrine granules with approx 100 nm size. SCLC are believed to derive from neuroendocrine cells of bronchial epithelium as closer morphological and genetic similarity are observed between large cell neuroendocrine carcinoma and SCLC (Hanna and Onaitis, 2013). SCLC are observed to spread along bronchi in a submucosal and circumferential fashion. SCLC consist of small cells with white-tan soft friable perihilar masses, huge nodal involvement, scant cytoplasm with ill-defined cell borders, finely granular chromatin and without or inconspicuous nucleoli. SCLC cells may be round, oval or spindle-shaped and necrosis is extensively observed in them as figure 5.10(B) shows. At times SCLC are found along with any of the histological subtypes may be ADCC, SQCC or LCC and may be called like adeno-squamous. SCLC are highly associated with smoking. SCLC are CD56 positive, have all neuroendocrine markers, TTF positive (Nicholson *et al.*, 2002). SCLC are aneuploid. Chromosomal imbalance occur as more than 90% cases of SCLC have large segments of 3p deletions and are associated with 3q isochromosome formation. Chromosomal subregions are amplified characteristically in SCLC. Normally SCLC tend to develop brain metastasis during progress of the disease. Paraneoplastic syndrome is common in SCLCs. They have high rate of p53 mutations. Myc amplification and methylation of caspase8, antiapoptotic gene are characteristics of SCLC. Over-expression of E2F1 and inactivation of retinoblastoma (RB) gene are unanimously found in SCLC (Nicholson *et al.*, 2002).

Adenocarcinoma

Adenocarcinomas are malignant epithelial tumor with glandular differentiation or mucin production characteristics. They typically express epithelial markers (AE1/AE3, CAM5.2, epithelial membrane antigen, Carcino-embryonic antigen, (CEA) (Colby, 1995). ADCCs show acinar, papillary bronchio-alveolar or solid with mucin growth pattern, with prominent nucleoli as shown in figure 5.10(C). Histologic heterogeneity is a characteristic of Adenocarcinomas and phenotypic expression is influenced by anatomic location. Central ADCC are supposed to arise from bronchiole epithelium or glands. The peripheral ADCCs consist of cells that are similar to type II pneumocytes or Clara cells. Peripheral

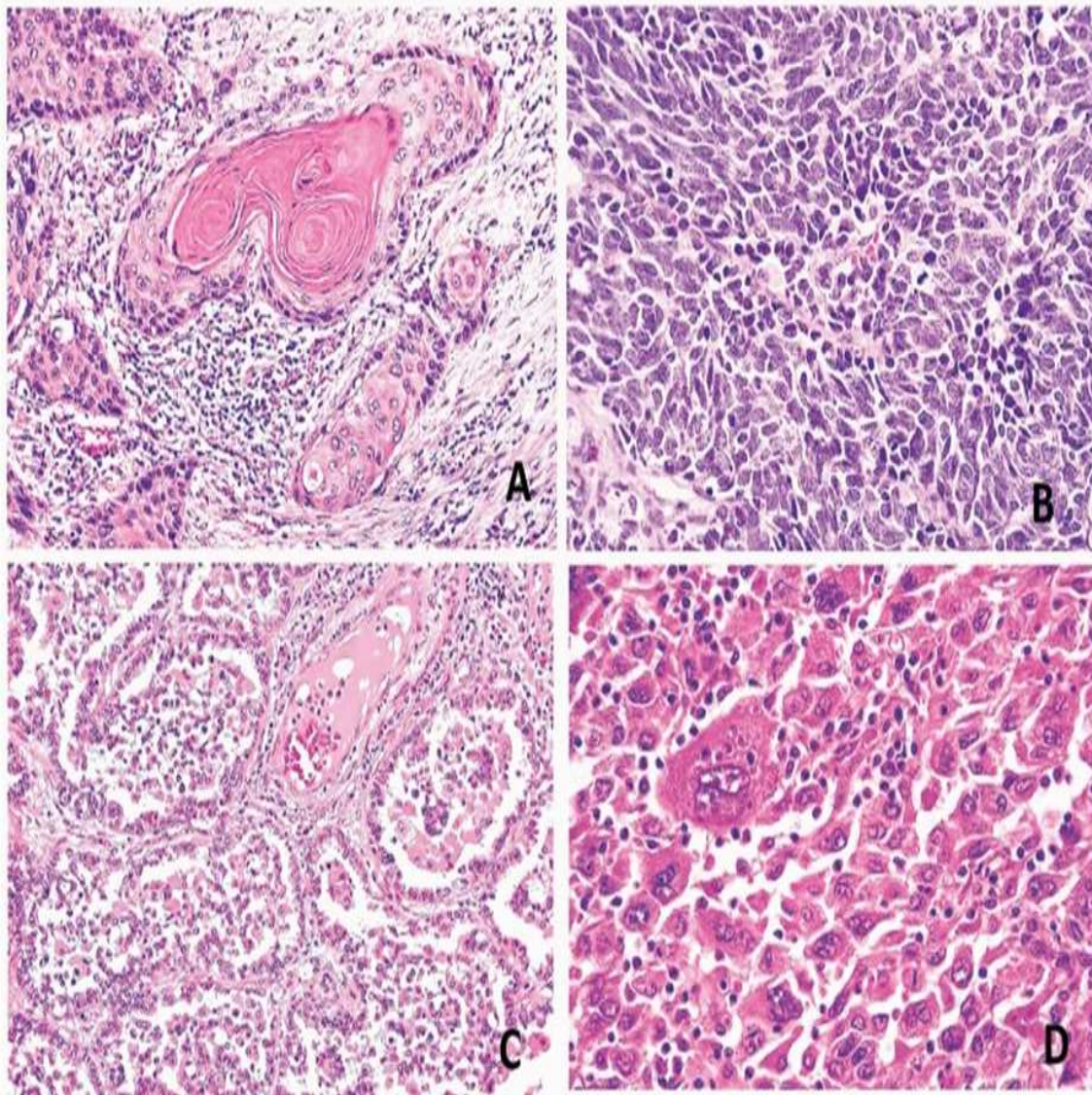


Figure 2.1: Microscopic images of different Histological subtypes of Lung Carcinoma (A) Squamous cell carcinoma (B) Small cell lung cancer (C) Adenocarcinoma (D) Large cell carcinoma (Source:(Weiss, 2005))

adenocarcinomas are grey-white central fibrosis with pleural puckering, edges of tumor may be lobulated or ill defined with stellate borders. Central or endobronchial tumor may have polypoid growth of tumor. Next pattern of tumor is diffuse pneumonia like, lobar consolidation occur. Fourth pattern gives an image of diffused bilateral lung disease involving all lobes appearance suggest diffuse pleural thickening. In fifth pattern tumor invades and extensively disseminates visceral pleura mimicking malignant mesothelioma. This histology is mostly found in non-smokers (Kondo *et al.*, 2002) and recently an alarming increase is seen in women. ADCCs are mostly found in peripheral nodules with rare cavitation. Recently ADCC is considered as the most common histology outnum-

bering SQCC. There occurs an over expression of COX-2 and HER2/NEU genes. CK7 is more expressed compared to CK 20. Characteristic immuno-phenotype of pulmonary adenocarcinoma tumor is that it stains positive for CK7 and negative for CK20 (Rubin *et al.*, 2001). TTF1 staining is positive in fully differentiated cells since some ADCC are mucin producing mucinous tumors figure 5.10. Adenocarcinomas are nearly diploid with loss of Y chromosome and gain of chromosome 1 and 7. Chromosomal imbalance occurs with 1q over-representation. Genetic alteration include point mutations of oncogene K-ras and TSG p53 and p16INK4a (Travis and Harris, 2004).

Large Cell Carcinoma

They are malignant epithelial tumor that are poorly differentiated. It accounts only 9% of lung cancer (Travis and Harris, 2004). LCC were previously called Large cell neuroendocrine carcinoma (Takei *et al.*, 2002). They are peripheral tumors and found in the periphery of lungs. Origin of LCC is believed to be bronchial preneoplastic lesion. Large cell carcinoma tumors are with soft pink tan with frequent necrosis occasional hemorrhage, and rare cavitation as indicated in figure 5.10(D). LCC consist of sheets or nest of large polygonal cells with vesicular nuclei and prominent nucleoli and moderate cytoplasm. LCC is mostly an aneuploid neoplasm with highest mean number of chromosome and DNA content of all lung cancer types. LCC show K-ras mutations, p53 mutations and Rb alteration like in other NSCLCs.

2.2.2 TNM classification

Staging of lung cancer is done based on Tumor-node-metastasis (TNM) classification given by 7th edition of TNM staging system by International Association for the study of lung cancer (IASLC) (Mirsadraee *et al.*, 2012; *Staging*, 2018). This includes the size and the degree of loco-regional invasions by primary tumor (T). The extent of lymph node involvement (N) and the absence or presence of intrathoracic or distant metastases (M). This classification system assists clinicians in defining prognosis for the patients and evaluate treatment results. Staging of lung cancer is a significant factor in determining the prognosis of the disease. The higher would be the stage, more extensive would be the spread of the disease, and less effective the prognosis (Rami-Porta *et al.*, 2009). Stage of lung tumor can be determined by two ways pathological and clinical. Clinical stage is determined as a result of imaging test, biopsies and physical test and pathological stage is determined after surgery in addition to the determination of clinical stage (Santos *et al.*, 2007). Once tumor is assigned TNM grading, stage can be concluded by analyzing the TNM (UyBico *et al.*, 2010; *American Cancer Society*, 2017). Table 2.1 summarizes

the classification for stage. Determination of the M status remains the primary target of clinicians, since any distant metastasis automatically leads to stage IV. This can be established with imaging tests only and no further analysis is needed. The tumor descriptors are subdivided based on the size and location which help to better define tumor classification. Tumor characteristics are determined by the same methods as diagnosing methods. Table 2.2 indicates the details of TNM classification and parameters used that help in staging.

Table 2.1: TNM stage grouping (Source:(Weissferdt, 2012))

Occult carcinoma	TX	N0	M0
Stage 0	Tis	N0	M0
Stage IA	T1a	N0	M0
	T1b	N0	M0
Stage IB	T2a	N0	M0
Stage IIA	T2b	N0	M0
	T1a	N1	M0
	T1b	N1	M0
	T2a	N1	M0
Stage IIB	T2b	N1	M0
	T3	N0	M0
Stage IIIA	T1a	N2	M0
	T1b	N2	M0
	T2a	N2	M0
	T2b	N2	M0
	T3	N1	M0
	T3	N2	M0
	T4	N0	M0
	T3	N1	M0
Stage IIIB	T1a	N3	M0
	T1b	N3	M0
	T2a	N3	M0
	T2b	N3	M0
	T3	N3	M0
	T4	N2	M0
	T4	N3	M0
Stage IV	Any T	Any N	M1a
	Any T	Any N	M1b

Table 2.2: Tumor node and metastasis Descriptors (Source:Rusch (2015))

—Primary tumor (T)—	
TX	Primary tumor cannot be assessed or tumor proven by the presence of malignant cell in sputum or bronchial washings but not visualized by imaging or bronchoscopy
T0	No evidence of primary tumor
Tis	Carcinoma in situ
T1	Tumor less than or equal to 3 cm in its greatest dimension, surrounded by lung or visceral pleura without bronchoscopic evidence of invasion more proximal than the lobar bronchus (that is, not in the main bronchus)
T1a	Tumor 2 cm or less in greatest dimension
T1b	Tumor more than 2 cm but 3 cm or less in greatest dimension
T2	Tumor with any of the following features of size or extent: ·greater than 3 cm in its greatest dimension ·may extend into the main bronchus if it remains more than 2 cm from the carina ·may invade the visceral pleura ·may be associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not involve the entire lung
T2a	Tumor more than 3 cm but 5 cm or less in greatest dimension
T2b	Tumor more than 5 cm but 7 cm or less in greatest dimension
T3	Tumor more than 7 cm or one that directly invades any of the following: parietal pleural (PL3), chest wall (including superior sulcus tumors), diaphragm, phrenic nerve, mediastinal pleura, parietal pericardium; or tumor in the main bronchus less than 2 cm distal to the carina but without involvement of the carina; or associated atelectasis or obstructive pneumonitis of the entire lung or separate tumor nodule(s) in the same lobe
T4	Tumor of any size that invades any of the following: mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, vertebral body, carina, separate tumor nodule(s) in a different ipsilateral lobe
—Regional lymph nodes (N)—	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastases
N1	Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes, including involvement by direct extension
N2	Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s)
N3	Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s)
—Distant Metastasis (M)—	
M0	No distant metastasis
M1	Distant metastasis
M1a	Separate tumor nodule(s) in a contralateral lobe, tumor with pleural nodules or malignant pleural (or pericardial) effusion
M1b	Distant metastasis (in extrathoracic organs)

2.2.3 Epidemiology of Lung cancer-Worldwide

Lung cancer is deemed as the most common cancer globally for last so many years. Different continents and countries around the world report their lung cancer statistics where lung cancer is becoming a pressing issue. The statistics record an overall increase in the incidence of lung cancer and mortality due to lung cancer. Smoking is considered as the major reason for the cause of lung cancer over the years. However recently a shift has been observed where the incidence rate has increased in females and non-smokers. The reason for this shift could be changing lifestyle, occupational exposure to radiations, increase in air pollution, passive smoking, increased consumption of liquor *etc.*.

WHO report for the year 2012 says that there were 1.4 (13%) million new cases and 19% deaths worldwide which were estimated to increase by 70% in near future. Year 2015 recorded 8.8 million deaths due to lung cancer which was 1 in every 6 deaths worldwide (*WHO*, 2018). Whereas there is an estimation of 222,500 new cancer cases being reported in 2017 (Travis and Harris, 2004). Lung cancer is found to be more prevalent in men worldwide (16.7%) of total compared to women. American cancer society in its surveillance for the year 2016 says 14% of patients in the disease are men (117,920) compared to 13% which are women (106,470). It reported 27% of males dead due to lung cancer compared to 26% females (*Cancer Facts & Figures*, 2016). Further, global rating as per Globacon says North America ranks first with the highest incidence rate of lung cancer with Africa ranking the last. Fact and data sheets show Hungary has the highest lung cancer risk with 51.6 rate per 100,000 (*Lung Cancer*, 2012). The regional lung cancer registry for Australia has ranked lung cancer as the 5th most commonly diagnosed cancer for the year 2013 and its estimated it would remain the same till 2017 with 12,434 new cases compared to 11,174 cases in 2013 (*Lung cancer statistics*, 2015). American cancer society accounted 1,685,200 new lung cancer cases and 595,690 deaths for the year 2016.

Cancer Research association in UK reported for the year 2014 that there were 46,403 new lung cancer cases with 35,895 deaths and 5% survival rate of 10 years and longer. Lung cancer was reported as the 3rd largest cancer in UK. Its projected that cancer incidence rates would fall by 7% between 2104-2035. There were 410,000 new cancer cases diagnosed in the year 2012 and 35900 deaths by the end of the year 2014 (*Lung cancer statistics UK*, 2015) . On the same lines whole Europe reported 354,000 people dead due to lung cancer. Similar to other data mortality rate was higher in males with 19,600 deaths compared to 16,300 deaths in females for the year 2014. However, for a change an improved 1-year and 5-year survival rate was observed for women.

2.2.4 Epidemiology of Lung cancer-India

During the early 19th century, the incidence of lung cancer in India was found to be extremely rare. In early 1950s 27.4 per million cases were reported and in 1959 it increased to 78.6 per million (Behera and Balamugesh, 2004). Adding to the above proceedings, Indian data also has analogous records. Recent Indian Council of medical research (ICMR) report accounts that nearly 17.3 lakh new cancer cases and 8.8 lakh deaths due to cancer would be registered in India by 2020 with lung cancer ranking second. Indian epidemiological surveys according to a report of ICMR for the year 2016 estimated 1.14 lakh new Lung cancer cases, where there were 83,000 males and 31,000 females. However, it was reported by the year 2020, the newly diagnosed lung cancer cases would increase to 1.4 lakh where 90% of them would be in advanced inoperable stage (*New cancer cases in India by 2020*, 2015). A consolidated report of Population based cancer registries under National cancer registry programme of Indian council of medical research stated in a review that the highest cancer incidence rate was observed in males in eastern region with 1 person out of 67 affected by lung cancer. The lowest rate of age adjusted cancer incidence was observed in Barshi a Western Indian rural cancer registry (Asthana *et al.*, 2015). Globacon 2012, revealed there were 63,759 deaths reported in India due to lung cancer making it the third largest reason for cancer death. Furthermore, the mortality was higher in males with 48,697 deaths compared to 15,062 dead females (Noronha *et al.*, 2016). Over a time of 24 years (1982-2005), ICMR has observed lung cancer cases per one lakh has increased by around 100% in Bangalore, 40% in Delhi and 160% in Chennai whereas a reduction of 60% was observed in Mumbai as figure 2.2 (*LUNG CANCER Focus:INDIA*, 2007).

High precedence of squamous cell type histologic variant was observed which is specific to smokers with adenocarcinoma ranking second. However recently a shift is being observed in histology as adenocarcinoma is surpassing the number of squamous cases in India along with other Asian countries. Adenocarcinoma is analyzed to be more prevalent in non-smokers. Increase in the number of non-smokers could be a reason for increase in adenocarcinomas (Malik and Raina, 2015).

2.2.5 Risk factors

Lifestyle constitutes the major risk factor responsible for the cause of lung cancer. The universal fact considered responsible for lung cancer is smoking. This could be the reason for alarming incidence and mortality due to lung cancer in men (Globacon, 2012). On the

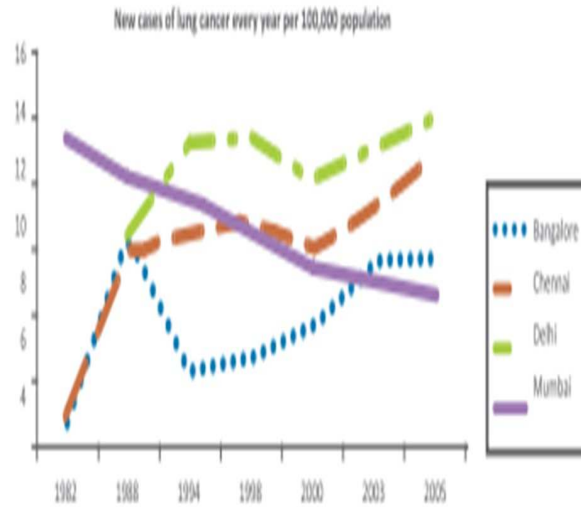


Figure 2.2: Lung cancer trend in four metro cities in 24 years (Source: (*LUNG CANCER Focus:INDIA*, 2007))

other hand no strong relations are prominently observed between smoking and females which indicate towards other risk factors like indoor air pollution, biomass fuel exposure, micronutrients insufficiency in diet *etc.* Surveys in UK reported 89% of lung cancer cases are due to smoking (*Lung cancer statistics UK*, 2015). Smokers to nonsmokers ratio in India is observed to be 20:1 which is though lower than western world. 80-90% cases in India are also attributed to smoking. Passive smoking, tobacco chewing, air pollution, diesel engine exhaust also add silently to the cause of lung cancer. Various risk factors are detailed below.

Tobacco Smoke

Universally tobacco smoke is considered as the major risk factor responsible for lung cancer. Lung cancer risk is higher in smokers than non-smokers. The severity of the risk increases the longer one smokes and the higher number of pack years one smoke. Worldwide it is observed that 90% of male lung cancer deaths and 75-80% female lung cancer deaths are caused by smoking (Ankathil, 2010). Epidemiological studies have forced us to conclude that tobacco smoke has carcinogenic effects (Hackshaw *et al.*, 1997; Sasco *et al.*, 2004; Fleiss and Gross, 1991). Geography and easy availability of tobacco reflect the pattern of tobacco consumption over the time (Peto *et al.*, 1992; Jemal *et al.*, 2010). Squamous cell carcinoma is majorally found in smokers. Recently a shift has been observed in histology from SQCC to ADCC and the use of newer low-yield cigarette is considered the reason for it as the site of disease has changed during time from bronchus or trachea to peripheral lung (Brooks *et al.*, 2005). People who are ex-smokers still have

a higher risk of lung cancer compared to non-smokers though there is a decrease in risk in ex-smokers compared to smokers.

Smoke from burning of tobacco contain nearly 1010 particles where 95% of it consists of gases chiefly nitrogen, oxygen, carbon dioxide, carbon monoxide, water, volatile sulphur compounds, acetaldehyde *etc.* (Hecht, 1999). There are nearly 3500 carcinogenic compounds in particulate phase of smoke. IARC (International Agency for Research on Cancer) has evaluated 55 carcinogens till date. Major particulate substances present are nicotine (0.2-0.6% of the weight of tobacco smoker), catechols (1%), polycyclic aromatic hydrocarbons (PAHs) (0.00002-0.00007%), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (Hoffmann *et al.*, 2001). In Indian subcontinent tobacco flakes rolled in a leaf of tendu plant called "bidis" forms the major form of tobacco consumption along with hookah and cigarettes. 73% of tobacco consumption is through bidis compared to 27% by cigarettes. Bidis are observed to have more nicotine than cigarettes (21.2 vs 13.5mg/g). A bidi smoker tend to make more number of puffs due to low combustibility of bidi leaf resulting in deleterious effect of bidis (Prasad *et al.*, 2010). Smoking index of an individual is used to calculate the cumulative risk due to smoking. Smoking status is quantified by "pack years". Pack years is defined as the product of average number of cigarettes smoked per day and the total duration of smoking in years (Jindal *et al.*, 1988).

Second hand smoke or passive smoking

Passive smoking is also considered equally harmful as there are reports where non-smokers show high risk of lung cancer (McCarthy *et al.*, 2012). Never smokers have reported 10-15% of all lung cancer cases. Side-stream smoke is reported to have carcinogens such as aromatic amines, PAH, Nitrosamines. Side stream has four times higher B[a]P than in main stream of smoke. Non-smokers exposed to second hand smoke are observed to have DNA adducts from tobacco carcinogens in urine samples (de Groot and Munden, 2012).

Diet and alcohol

The recent fast pacing life with change in our food habits have caused a damage to our health. The dietary habbits have changed over the time from home made food to packaged food with heavy spices and oils. It is documented that diet rich in vegetables, fruits, cruciferous vegetables have high iso-thiocyanates present in them, which could be responsible for reduce risk of lung cancer (Lam *et al.*, 2009). Smokers with high alcohol consumption are observed to be at high risk of lung cancer (Malhotra *et al.*, 2016).

Occupational exposure

Aetiology of lung cancer is undilutedly influenced by occupational exposure:

Asbestos They are tiny hair like fibres that are found in rocks. Asbestos has a fire resistant quality due to which it is tempted to be used in construction such as factory work, insulation, carpentry, demolition. Asbestos fibre become lodged in the lining of the lung and is observed to be responsible for mesothelioma type of lung cancer (Berman and Crump, 2008).

Radon Countries report radon as the second most important cause of lung cancer and the leading cause of lung cancer in non-smokers (Gawelek *et al.*, 2017). Radon is odour less gas without smell. It is produced by soil and rocks that have uranium. The gas from them can travel into air and water. Its said basements have higher level of radon as it is near to earth. Radon gas produces tiny radioactive particles that can attack the lung lining and produce radiations which can cause ultimately lung cancer.

Metals and mixed occupation exposures

People working on Chromium (IV) compounds, nickle miners have reported an increased risk of lung cancer (Hayes, 1997). High risk of lung cancer is also observed in people exposed to arsenic may be in drinking water (on the Evaluation of Carcinogenic Risks to Humans *et al.*, 2004). People who are in occupation related to silica such as poultry, brick making, stone cutting, earth mining, have also reported an increased risk of lung cancer, with evidence of exposure-response relationship in the high-exposure range (Steenland *et al.*, 2001).

PAH and Diesel exhaust and air pollution

Combustion of organic material produces polycyclic aromatic hydrocarbons (PAH). Aluminium, coal gasification, coke production, roofing, tar distillation produces PAH and an increased risk of lung cancer has been reported in people exposed to high PAH (Rota *et al.*, 2014). Motor vehicle engine exhaust represents a high risk causative factor for lung cancer. Lung cancer is immensely observed in non-smoking women, researchers attribute indoor air pollution as the reason for the disease in them. Coal burning in non-ventilated houses, working on chullahs, fumes from high temperature cooking using unrefined vegetable oil (Melville and Eastwood, 1998).

2.3 Lung cancer is Age/Gender based

Traditionally lung cancer is defined as the disease of males and adults. However, the scenario has changed over the time as women have also opted to smoke and continuous industrialisation has made the air so polluted. The growing evidences suggest that the lung cancer irrespective of its subtype and no of cigarettes smoked has three times higher

risk in women than men (Papadopoulos *et al.*, 2014; Risch *et al.*, 1993; Osann *et al.*, 1993). Non-smoker females also show higher lung cancer risk with adenocarcinoma being the most common histology (Gasperino, 2011). Researches have hypothesized that estrogen metabolism and exposure to tobacco smoke constitute the major factors for higher lung cancer carcinogenesis in females (Tansavatdi *et al.*, 2004). PAHs in smoke activates CYP1B1 gene which metabolizes 17- β estradiol (estrogen), female hormone. Estrogen is present in two forms ER α and ER β . ER β is a substrate for CYP1B1 and both are found in the lung epithelium. Genotoxic catechol estrogens are formed after the enzymatic metabolism of E2. The catechol estrogens bind to DNA and form adducts in genes like tumor suppressor genes, or proto-oncogenes (Gasperino, 2011). In addition to this it is revealed that women have lower DNA repair capacity therefore lower removal of nicotine from plasma (Rivera and Stover, 2004). Smoking females with lung cancer were observed to have elevated levels of CYP1A. The level of adduct formed guide the carcinogenesis, higher the level of CYP1A higher is the adduct formed, which was subsequently found higher in females (Wei *et al.*, 2000). Further risk for lung cancer is observed to increase with age both in smokers and never smokers. It is suggested mostly lung cancer patients are above 65 years of age (Tas *et al.*, 2013; Brown *et al.*, 1996). The lung cancer incidence rate and mortality both are observed to be higher in patients aged >65 (Yancik, 1997). Hereditary nature of lung cancer is not well known, however to some extent the first degree family member with lung cancer increases the risk level towards lung cancer (Coté *et al.*, 2012). Though Lung cancer is considered a disease of acquired mutations during lifetime, inheritance of certain genetic modification put people at higher risk of lung cancer irrespective of smoking for instance people with EGFR mutation have higher chances of familial lung cancer (Gaughan *et al.*, 2013).

2.4 Lung cancer in Never smokers

The term "never smoker" is given to a person who smokes lesser than 100 cigarettes during lifetime. Lung cancer is traditionally an outcome of smoking however, digging of prior data indicates that 10-25% of lung cancer occurs in non-smokers worldwide (Couraud *et al.*, 2012). However, across continents it was revealed, Southern east Asia reported 50% women with lung cancer and they were non-smokers. This record decreased to 2-6% in men in Western areas (Couraud *et al.*, 2012). As discussed above there are many causative factors for lung cancer except smoking *i.e.* second hand smoke, radon exposure, occupational exposure, indoor air pollution specifically in women, air pollution, diet etc (Samet *et al.*, 2009). Most of the cases of lung cancer in non-smokers

are adenocarcinomas. The gene profiling of non-smokers exhibiting lung cancer is different from that of lung cancer. It is observed that non-smokers have high expression of EGFR and tyrosine kinase mutations. The genes K-RAS and P53 exhibit different mutation profile for non-smokers than smokers (Subramanian and Govindan, 2008).

2.4.1 Lung Carcinogenesis

Lung cancer is a paradigm of carcinogenic induced cancer in constricted association with genetic predisposition. The on-set of lung cancer occurs when normal epithelial cells lining the respiratory tract are exposed to carcinogens in smoke over a period of time (Cruz *et al.*, 2011). The carcinogens present in tobacco smoke are first metabolically activated and sequentially excreted from the body (Control and Prevention (2010); Hecht (1999)). During the process DNA-adducts are formed (Phillips, 2002). Cellular repair system removes the DNA adducts and renders the normal DNA. However, when DNA damages exceed the DNA repair capacity, DNA damages persist and lead to the formation of mutations. Oncogenes and Tumor suppressor genes if become the target of these mutations, they tend to lead towards unregulated growth and ultimately carcinogenesis. Figure 2.3 gives a framework for lung carcinogenesis.

Metabolic activation and excretion of Smoke carcinogens

The carcinogens present in smoke are substrates for drug metabolic system. The enzymes that act upon on smoke carcinogens are cytochromes P450 (Phase I), Glutathione-S transferases and UDP-glucuronosyl transferases (Phase II) (Hecht, 1999). They sequentially act on the carcinogenic substrates and ultimately excrete it out of the body. The first step is activation of carcinogenic compounds which is done by cytochrome P450 (Hecht, 2012). The CYPs do either of the function given N- and o-dealkylation, aliphatic and aromatic hydroxylation, N- and S-oxidation or deamination (Taxak and Bharatam, 2014). The activated carcinogens formed during the process can actively bind with DNA and form DNA adducts. The by-products of Phase I, if polar enough, they are directly excreted out of the body, however otherwise Phase I is followed by conjugation of a polar compound to the Phase I byproduct. The Phase II conjugation is done by "transferase" enzymes". The process is done by Glutathione S-transferases (GSTs), UDP-glucuronosyl transferases (UGTs), Sulfatases (STs), arylamine N-acetyl transferases (NATs) and methyl-transferases which convert activated carcinogens in harmless more hydrophilic compounds that are easily excreted out of the body (Control and Prevention, 2010). The GSTs perform cytosolic dimerization, UGTs do glucuronidation, STs are responsible for sulfation of aliphatic alcohols and phenols. Phase II detoxification is followed by Phase III

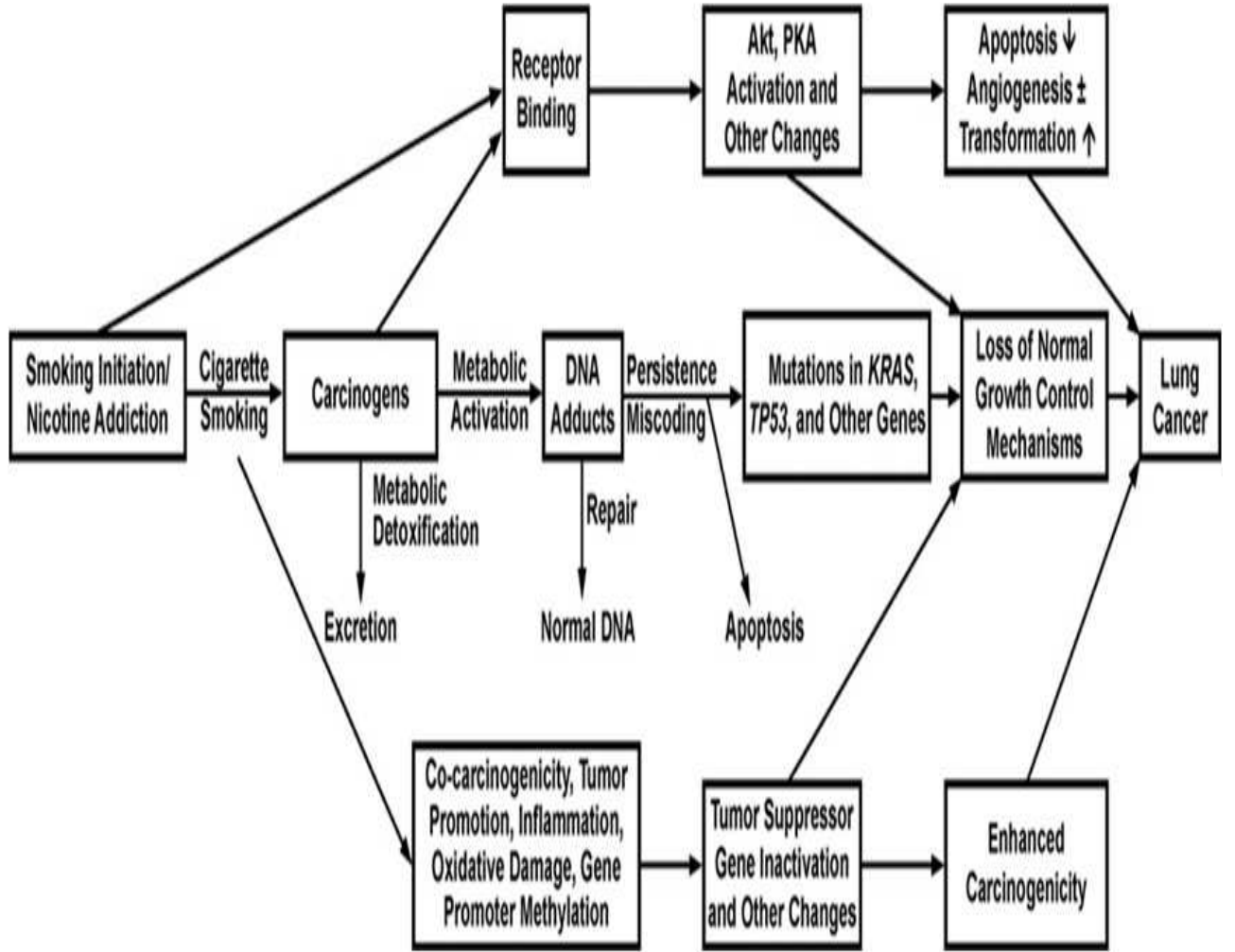


Figure 2.3: Framework for lung carcinogenesis

where membrane transporters come into activity. ABC transporters are important gene that form Phase III (Omiecinski *et al.*, 2010). There should exist a balance between the activation and detoxification of carcinogens. There occurs an inter-individual difference in the balance between activation and detoxification of carcinogens that ultimately affects the cancer susceptibility. Certain carcinogens are observed to directly form adducts without activation (Cohen and Arnold, 2010).

2.4.2 Molecular alterations in Lung carcinogenesis

Molecular alterations as DNA adducts were observed in respiratory epithelium (Kadara *et al.*, 2016). These DNA-adducts are either repaired or cell may lead to apoptosis (Roos and Kaina, 2006). In case cell is neither repaired nor apoptosized genetic instability oc-

curs through acquired genetic and epigenetic mechanisms in cells leading to mutational events and ultimately cancer. Whole chromosome gain or loss, DNA segment amplification or deletion, gene silencing through methylation all these methods leads to genetic alteration. These alterations are evident in karyotype and molecular profiles of a tumor. Mutations giving rise to lung carcinogenesis follow a sequence. Molecular changes with allelic loss in 3p regions such as 3p21, 3p22-24 etc. and 9p21 and to some extent in 8p21-23, 13q14, 17p13. A certain sequence is observed for allelic loss 3p-8p-9p (Wistuba *et al.*, 2002). Pathogenesis of lung cancer results after continuous accumulation of genetic alteration over the time which involves activation of oncogenes, inactivation of certain tumor suppressor genes, allelic losses, loss of heterozygosity (LOH), chromosomal imbalance as indicated in figure 2.4 .

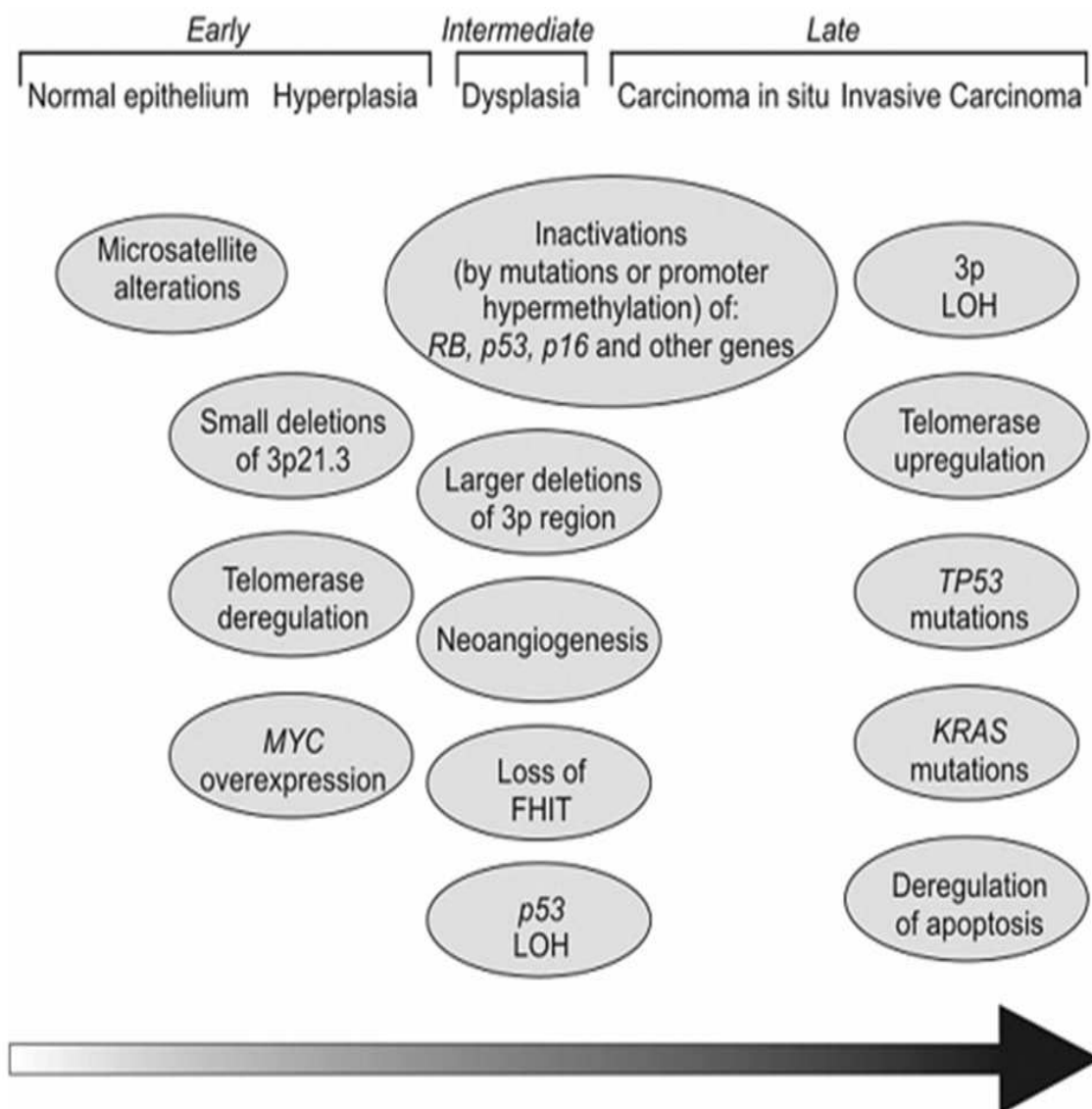


Figure 2.4: Molecular changes during lung cancer pathogenesis (Source:(Panov, 2005))

Activation of Proto-oncogenes

Tumor cells show production of growth factors for their autocrine and paracrine growth. Cells that produce growth factors and express its receptors show autocrine growth. Paracrine growth occurs when growth factors stimulate growth in adjacent cells. ERBB family is a group of receptor tyrosine kinases which is responsible for growth in case of NSCLCs (Kirschbaum and Yarden, 2000).

Epidermal growth factor receptor (EGFR) is responsible for events such as cell division, motility, neo-vascularization by ligand activated cascade of auto-phosphorylation of mediators and finally effector molecules. Signal transduction stimulated by EGFR occurs through PIK3/AKT/mTOR, RAS/RAF/MAPK and JAK/STAT pathways. Tyrosine kinase is constitutively expressed due to mutations occurring in EGFR gene. Mutated EGFR is considered responsible for oncogenic differentiation, proliferation, neo-vascularisation, invasion, metastasis *etc.*. Studies have reported that NSCLC subjects harbor most common mutations in four exons namely 18-21. Mutation in exon 19 is the most common one which is frame deletion mutation. Second most common mutation is in exon 21, this is a missense substitution (Kosaka *et al.*, 2004). EGFR mutations are mostly reported in adenocarcinoma subjects (Cooper *et al.*, 2013).

Tyrosine kinases inhibitors are extensively explored for their role in lung cancer since tyrosine kinases (growth related protein kinases) act as a therapeutic target for cancer. Similarly, ATP-binding domain of tyrosine kinases behave as a target for drug development (Thomas *et al.*, 2012). Gefitinib and erlotinib are two TKIs which are observed to have improved survival in lung cancer subjects (Ansari *et al.*, 2009).

HER2/neu (Human epidermal growth factor 2) and **c-myc**. HER2/neu is another member of ERBB family. It is a transmembrane glycoprotein of 185 kDa. Exon 20 of the gene shows activating mutation. Its expressed in NSCLCs and majorly in adenocarcinomas (Fong *et al.*, 2003). It was first observed that c-myc was over-expressed in 80-90% SCLCs (Rom *et al.*, 2000). The position of these phospho-proteins is nucleus and it belongs to transcription factors family. It also interacts with the histone deacetylase which directly governs the chromatin conformation responsible for affecting gene transcription (Forgacs *et al.*, 2001).

KRAS is a important member of RAS family of proto-oncogenes and encodes an important G-protein. It plays an important role in signalling pathways leading to differentiation, survival, and cell proliferation. In normal state Ras is inactive and GDP (guanosine diphosphate) bound. It gets activated when GDP switches to GTP and activates upstream growth factor receptors. The activated Ras-GTP consequently binds and activates MAPK, RAS/RAF/MEK/MAPK (Goodsell, 1999). The above signal cascade

plays a role in lung cancer with a mutation in K-RAS (Downward, 2003). Following K-ras mutation lung epithelial cells can transform by activating ERK-MAP pathways (Massion and Carbone, 2003). Mutated K-ras was observed in 15-20% NSCLC in adenocarcinomas. Most common mutation in K-ras occurs in codon 12, 13 and 61 (Riely *et al.*, 2008). K-ras mutation is reported in smokers with G-T transversion.

Tumor suppressor genes

(TSGs) or recessive oncogenes inactivated in the process include Tp53(17p13), RB (13q14), and p16INK4a (9p21). There are certain new tumor suppressor genes at short arm of chromosome 3 3p21(DTT1), 3p14.2(FHIT), 3p21(RASFF1), 3p22-24 (BAP-1). Inactivation of recessive oncogenes is a two step process, primarily point mutation occurs in first allele followed by inactivation of second allele may be by translocation or may be promoter methylation or chromosomal deletion (Wistuba *et al.*, 2002).

Tp53 is located on 17p13 chromosome and encodes for a protein of 53kDa and behaves as a transcription factor (Rivlin *et al.*, 2011). Carcinogenic load or damaged DNA induces p53 to dys-regulate cell-cycle and induces apoptosis and ultimately tumor formation (Zilfou and Lowe, 2009). Mutation in TSG **p53** are most commonly observed in SQCC and SCLC subtype of lung cancer (Wistuba *et al.*, 2002). p53(17p13) is likely to be used as a biomarker as over two-third of lung cancer report mutated p53. NSCLCs have reported *Tp53* mutations to be associated with a smoking history of patients (ShollLM, 2016). There are reports where mutations in *Tp53* is correlated with resistance towards lung cancer treatment. Next important tumor suppressor gene involved in lung carcinogenesis is **Rb(13q14)**. Rb cooperates p53 in transcriptional level, cell cycle progression and equilibrium between cell differentiation and proliferation. Non-sense mutations, chromosomal deletions are normal gene behaviour observed in lung cancer (Forgacs *et al.*, 2001). It is observed to be over-expressed in 90% SCLC and 20% NSCL cases (ShollLM, 2016). Rb acts as a control for cell to go from G1 to S phase of cell cycle. **p16 (9p21)** is another TSG. It is also termed as CDKN2A. Over 40% NSCLC report inactivated p16. Inactivation may occur by loss of heterogeneity, point mutation .

2.4.3 Diagnosis of Lung cancer

The diagnosis of lung cancer is finalized after a series of tests. The sequence of the tests starts with physical examination, medical history *etc.* followed by complete blood count, hepatic transaminase, calcium levels and measurement of alkaline phosphatase and chest radiography. Confirmation of tumor is done by imaging tests by Computed tomography scan (CT-scan), Positron emission tomography (PET-scan), Magnetic resonance imaging

(MRI), bone scan depending on the severity of the disease respectively.

The step followed by imaging test is confirmation of histology using techniques immuno-histochemical characterization. In order to carry out this molecular testing, a section of the lung tumor tissue is required. The approach used to obtain the tissue largely depends on the location and size of the tumor and its metastatic accessibilities. For the centrally located lesions, fibre optic bronchoscopy is used, whereas for lesions located near the periphery CT-guided transthoracic needle aspiration is a typical method. Endobronchial ultrasound and electromagnetic navigation are the techniques of the new diagnostics (Latimer and Mott, 2015)

2.4.4 Lung cancer Treatment

Treatment for lung cancer depends on stage of lung cancer and type of lung cancer. It's also required that extent of spread of lung cancer must be known before administering the treatment for cancer. Treatment for lung cancer includes surgery of tumor, chemotherapy or radiotherapy (RadiologyInfo.org, 2017). When tumor has not spread to distant areas and is limited to small area, it is treated with local treatment such as radiotherapy and surgery. Chemotherapy and targeted therapy are preferred when tumor has spread to distant areas, as in stage III B and IV. At times chemotherapy is given before or after local treatment to remove the remaining tumor cells (Chang, 2011). Chemotherapy can be given either before surgery as neo-adjuvant to shrink cancerous cells or along with radiotherapy as concurrent therapy where cancer has metastasizes to nearby important organs due to which there is no possibility of surgery or after surgery as adjuvant chemotherapy to kill any cancer cells that are left behind or as main treatment. When chemotherapy is given as main therapy cisplatin/carboplatin along with either paclitaxel, docetaxel, gemcitabine or vinorelbine is given in case of NSCLC. SCLC subjects are given cisplatin/carboplatin along with irinotecan (Behera, 2007).

Targeted therapy unlike chemotherapy targets the molecules expressed on the surface of cancerous cells. The therapy attack these targets either by blocking or destroying the targets. Anti-angiogenesis therapy is used to limit the supply of blood to tumor to avoid its growth and ultimately starve it to death *i.e.* anti-VEGF therapy (Al-Abd *et al.*, 2017). Vascular endothelial growth factor (VEGF) is responsible for angiogenesis in tumor. VEGF promotes survival of cells and further inhibit their apoptosis. Its a growth factor that even regulates the permeability of Vascular endothelial cells. Cancerous cells produce vascular endothelial growth factor. Anti-angiogenic compounds are hypothesized to act either by making the blood flow normal within the tumor and thereby improving

the efficacy of the treatment. Bevacizumab is FDA-approved VEGF inhibitor used in colon cancer. EGFR inhibitors eg. Erlotinib, Gefitinib, Afatinib block the growth of cancerous cells that are EGFR positive. EGFR provide for the development of cancerous cells growth and development (Cabanero *et al.*, 2017). Matrix Metalloprotease (MMPs) inhibitors: MMPs cause the breakdown of structure present between cells to allow for the growth of healthy new tissues. These MMPs are therefore responsible for the blood vessel development tissue growth and wound healing in normal cells. Cancerous cells invade MMPs for their growth and metastasize to other tissues and organs of the body.

2.5 DNA damage and Repair

DNA is a repository of genetic information present in a cell and it is under constant interaction of noxious agents which are DNA damaging. DNA damages may interfere with the steady state quasi equilibrium and give way to certain biochemical pathways leading to damage removal since genomic integrity is to be maintained (Clancy, 2008). DNA damages may be generated from exogenous sources such as environment or from endogenous source spontaneously at the time of DNA metabolism (Tubbs and Nussen-zweig, 2017). Its estimated spontaneous base loss in DNA is 10^4 base per cell per day resulting in development of abasic sites (Martin, 2008). Impulsive DNA alteration can be the result of mis-incorporation of dNTPs during replication, deamination, alkylation, depurination (Ciccia and Elledge, 2010). Additionally, metabolic process such as respiration, cell injury, hydroxylation, inflammation causes the generation of reactive oxygen species (ROS) (Birben *et al.*, 2012). ROS cause the modification and release of purine or pyrimidines eg. 5-hydroxycytidine, and 8-hydroxyguanine (Cadet and Wagner, 2013). This modification primarily occurs in GC bases compared to AT bases. Some of the modified bases have mutagenic properties which if not repaired properly leads to carcinogenesis, or some of them constitute block to replication while others are able to mispair with DNA and get incorporated in DNA (Cadet and Wagner, 2013).

Exogenous damaging sources can be broadly classified into physical or chemical. Physical factors are UV-radiations (200-300nm) from the sun, ionizing radiations *etc.* UV is responsible for pyrimidine dimer formation eg. Cyclobutane pyrimidine dimers (CPD) and 6-4 photo-products (6-4PPs) and they are responsible to distort the DNA helix (Budden and Bowden, 2013). Further, ionizing radiations may be X-rays, gamma rays, α particles, β particles *etc* when come in contact with DNA in a cell they tend to either directly pass the cell without causing any harm or physically damage DNA either in direct or indirect way. In case of direct damage the ionizing radiations directly interact with

atomic electrons of DNA through columbic forces. However, in case of indirect mode of action the radiations interact with water causing the generation of free radicals, which are highly reactive and can attack DNA (Desouky *et al.*, 2015). α particle when come in contact with DNA, it damages the sugar-phosphate backbone of DNA. The break in sugar-phosphate backbone can be in single strand causing single strand break or in both the strands of DNA causing double strand break (Reisz *et al.*, 2014). Single strand break can be repaired easily as the complementary strand is present as template to repair the damage caused. However, double strand breaks become detrimental as they can cause loss of DNA fragments and further leading to the joining of non-homologous chromosomes. β particles being of low intensity tend to break hydrogen bonds between bases and further X-rays cause damage to bases creating abasic sites (Curiosity, 2013).

Chemical mutagens are mostly alkylating agents and azides. They create a huge variety of DNA adducts such as ethenobases, oxidized bases, alkylated phosphotriesters etc. Alkylating agents add alkyl group to bases and may change the normal base pairing eg. alkylating agent EMS (Ethylmethanesulfate) convert guanine to 7-ethyl guanine which pairs with thymine (Griffiths AJF, 2000). Chemical agents like nitrous acid cause deamination of bases resulting in conversion of cytosine to uracil adenine to hypoxanthine and guanine to xanthine (Caulfield *et al.*, 1998). This deamination affect the hydrogen-bonding potential finally leading to mis-pairing. DNA damage is counteracted by DNA damage response where four responses *i.e.* DNA repair, transcriptional response, DNA damage checkpoints and apoptosis work sequentially (De Zio *et al.*, 2013). As a response to pernicious agents cell cycle check points come into activity and damage is either corrected or otherwise cell is apoptosized (Dasika *et al.*, 1999). Rectification is carried out by DNA repair machinery of the cell following the repair pathway based on the type of lesions as shown in figure 2.5. There are pathways like

- Direct protein reversal which act on lesions like O6 alkylguanine.
- Mismatch repair which act on DNA mismatches and insertion/ deletion loops arising from DNA replication.
- Base excision repair which act on abnormal DNA bases, base adducts.
- Nucleotide excision repair which act on heavy base adducts and UV photo products.
- Double strand break repair which act on double strands break damages.

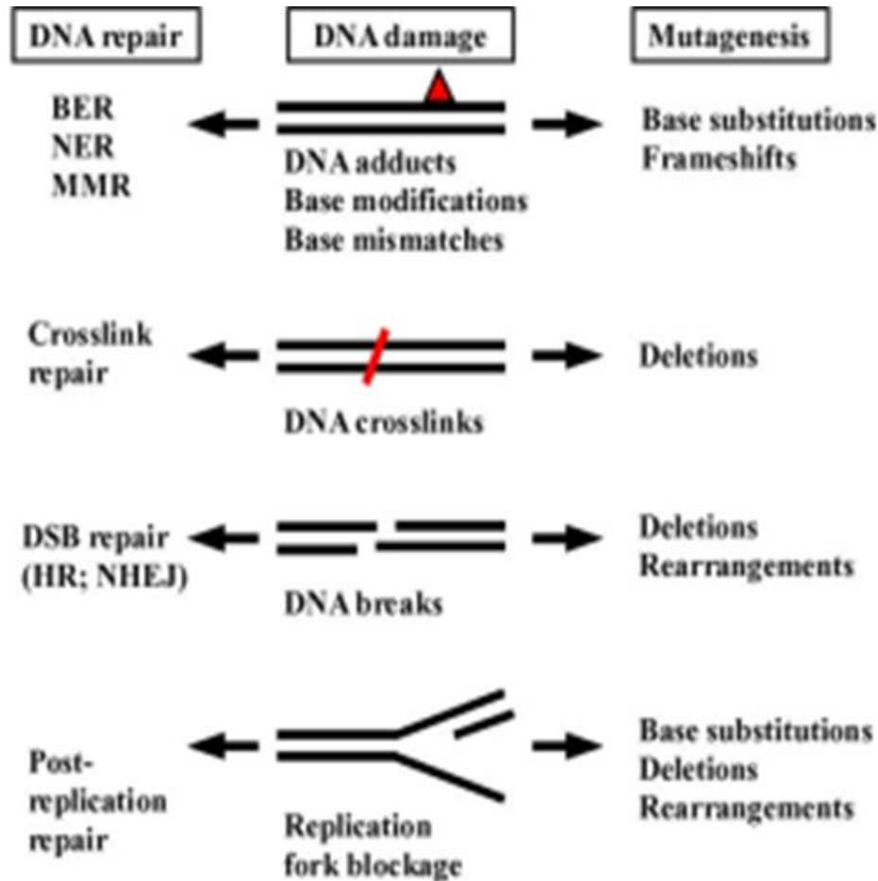


Figure 2.5: Specific DNA damage are responded by specific pathways (Source:(Dixon and Kopras, 2004))

2.5.1 DNA damage Repair pathways

Base excision repair Pathway(BER)

Base Excision repair is a highly coordinated enzymatic pathway and its activity comes into play when DNA encounters damage either by oxidation, methylation or from deamination (Krokan and Bjørås, 2013). The sites of BER is nuclei or mitochondria and whenever a damage is encountered DNA glycosylases come into activity and cleave the glycosidic bond between base and sugar ultimately generating an abasic site as shown in figure 2.6. Till date 12 DNA glycosylases are found which work depending on the base damaged. Consecutively the next enzyme in the sequence is Apurinic or apyrimidinic endonuclease I (APEI), which act on the abasic site generated (Kim and M Wilson III, 2012). The phospho-diester backbone becomes the target of APE1, which hydrolyzes the backbone 5' to the abasic site giving rise to single strand break harbouring 3'OH and 5'deoxyribose phosphate (5'drp) termini (Krokan and Bjørås, 2013). DNA glycosylases are observed to have an AP lyase activity along with base cleaving ability and via β -elimination generate

5' phosphate and 3'-phospho- α , β unsaturated aldehyde on the backbone. Subsequently NEIL1 (nei endonuclease VIII 1) and NEIL2 (nei endonuclease VIII 2) a subset of DNA glycosylases cause β and δ elimination changing the 3' blocking agent to 3'phosphate. Immediately, after an incision at the DNA backbone is made blocking lesions are recruited to the termini. However, for complete repair process to occur the DNA termini are to be restored to the normal 3'OH and 5'phosphate ends (Simonelli *et al.*, 2005). APE1 has characteristic additional intrinsic 3'-phosphodiesterase property which aids in restoring 3'OH from blockades 3'-phospho- α , β unsaturated aldehyde. On the other hand PNKP (Polynucleotide kinase 3'-phosphatase) helps to convert the 3'phosphatase to 3'OH and 5'dRP to 5'-phosphate by dRP lyase activity of DNA pol β (Beard *et al.*, 2006). Single strand breaks mostly have end blocking agents at ends, therefore besides canonical enzymes, BER pathway has enzymes that work specifically for restoring DNA ends. 3'phosphoglycolate and 3'phosphoglycoaldehyde are the frequent blocking ends developed after ROS induced breaks (Hegde *et al.*, 2011). 3'-phosphodiesterase property of APE1 again acts here to restore the ends. There are other end processing enzymes like Tyrosyl-DNA phosphodiesterase (TDP1) and Aprataxin (APTX) which acts on 3'phosphoglycolate and 5'adenylate intermediates respectively. Once the required 3'OH and 5'phosphate ends are regained the next sequential step *i.e.* repair of DNA occurs followed by ligation (Hegde *et al.*, 2011).

The damage to be repaired may be single base or number of bases, this classifies the process into either *long patch* repair or *short patch* repair. In case of short patch repair 5'dRP gets removed by lyase activity of ploy β and final ligation occurs either by ligase I or a complex of ligase III and *XRCC1* (Dexheimer, 2013). Approx 80-90% of repair in BER is through short patch repair. Repair occurs as long patch when lyase activity of DNA polymerase β is not able to act upon 5'blocking lesions (Almeida and Sobol, 2007). DNA polymerase β , δ , ϵ along with PCNA extends 3'OH into gap and remove the 5'-lesion as a flap. FEN1 (Flap endonuclease-1) comes and then removes the flap that is generated. This step is then followed by the action of ligase I that is recruited and ligates the gap which will move downwards from the original damage site.

Genes *OGG1*, *MUTYH* and *XRCC1*: Glycosylases are responsible for recognizing specific damaged bases and removing them. Gene 8-Oxoguanine glycosylase *OGG1* codes for DNA glycosylases that specifically cleave N-glycosidic bond between deoxyribose and 8-oxoguanine (8-OxoG) (Faucher *et al.*, 2012). Oxidative damage to guanosine results in the formation of 8-Oxo-G which can base pair with adenine and cytosine. Uracil DNA glycosylases can cleave specifically uracil. However, thymine-DNA glycosylase can cleave Uracil and thymine both when they are mis-paired with guanosine (Faucher *et al.*, 2012).

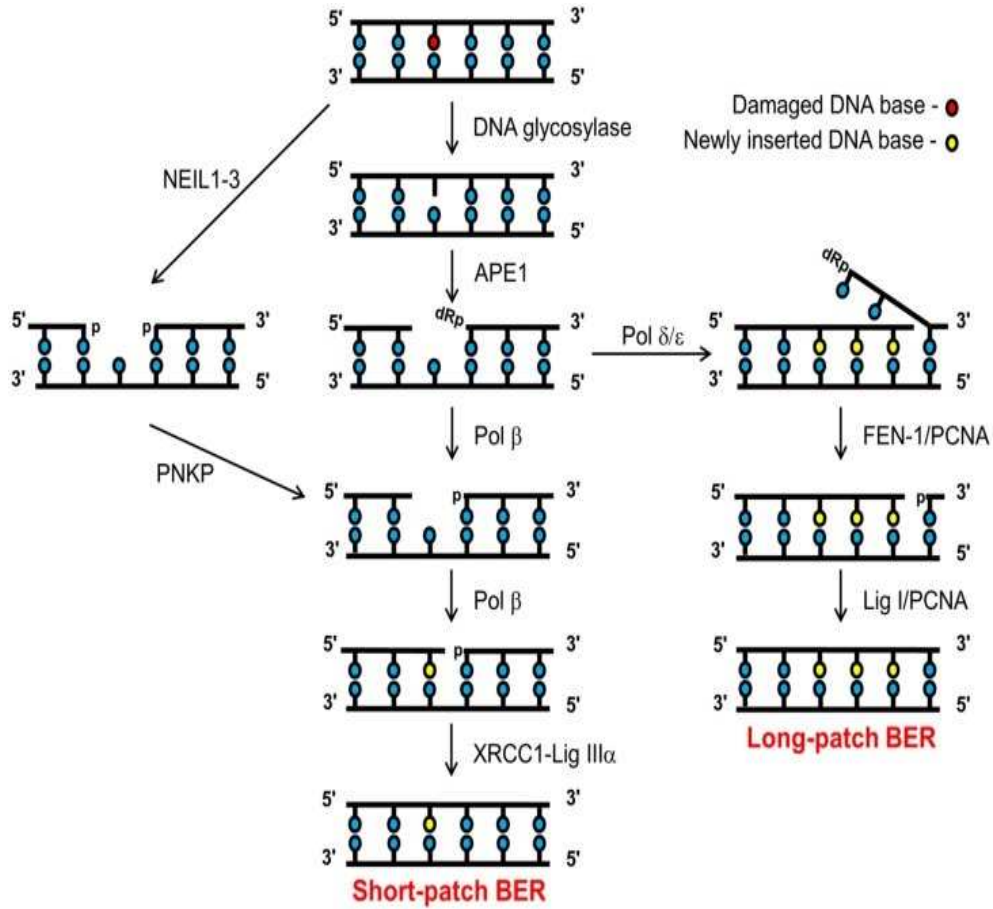


Figure 2.6: Base excision repair pathway

The gene *OGG1* is located on chromosome 3 at position 3p25 and it codes for protein glycosylase 8-oxoguanine (Waters *et al.*, 1999). The gene has seven exons and codes for a protein of 345 amino acids and the protein coded by *OGG1* is bi-functional in nature (Fortini *et al.*, 1999). The unique feature of *OGG1* protein is that it has alternate mode of splicing of the C-terminal region. α *OGG1* is a protein of 345 amino acid residues and β *OGG1* is a 424 amino acid protein (van der Kemp *et al.*, 2004). N-terminal remains the same for the two proteins however the difference comes due to the last exon of the gene (Sameer *et al.*, 2012).

(mutY Homolog (E.coli)) MUTYH is another DNA glycosylase or counterpart of *OGG1* in the chain which removes the adenine mispaired with 8-oxoG. The gene is located on chromosome 1(1p34.1) at short arm and spans a total of 11.2kb region (Oka *et al.*, 2014). The protein for the gene has a domain based structure and the N-terminal region has catalytic area with helix-hairpin-helix(HhH), iron-sulphur cluster loop motif and pseudo HhH (Kashfi *et al.*, 2013).

XRCC1 (X-ray cross complementing protein I), is a structural protein plays a very impor-

tant role in BER pathway. *XRCC1* acts as a scaffold protein for the DNA assembly polymerase, DNA ligase, PNKP (Polynucleotide kinase 3'-phosphatase), APTX (aprataxin), TDP1 (Tyrosyl-DNA phosphodiesterase 1) and the nicked DNA to be repaired Katyal and McKinnon (2011). *XRCC1* interacts with DNA polymerase β at NH2 terminal and with DNA ligase III at COOH terminal (Kubota *et al.*, 1996). *XRCC1* is the gene that affect the sensitivity towards ionizing radiations. The gene *XRCC1* maps to chromosome 19 at short arm q and spans a region of 31.9kb. It has 17 exons and codes for a protein of 633 amino acids with a molecular weight of 70KD (Sterpone and Cozzi, 2010). It is tightly associated with ligase III since *XRCC1* is essentially required for the stability of Ligase III (Lakshmipathy and Campbell, 2000). Studies on EM9 and EM11 with *XRCC1* mutant have shown that levels of Ligase III are reduced (Caldecott *et al.*, 1996).

Double strand break repair pathway (DSBR)

Double strands breaks (DSB) are a serious threat to DNA as they can affect the viability of cells. Double strands breaks are the result of DNA exposed to ionizing radiations, chemical agents or UV lights (Rastogi *et al.*, 2010). They can be naturally produced in cell during V(D)J recombination, class switch recombination, meiosis, during mitosis when DNA is pulled towards opposite poles. If double strand break is unrepaired it may lead to translocations, deletions, chromosome fusions (Varga and Aplan, 2005). Eukaryotic double strand break repair occurs by two pathways Homologous recombination (HR) and Non-homologous end joining pathway (NHEJ) (Shrivastav *et al.*, 2008).

- Homologous Recombination- Homologous recombination pathway is termed as "error-free" pathway. HR requires sequence homology to serve as a template for the repair of broken ends (Meers *et al.*, 2016). HR can be divided into three major phases: Presynapsis, synapsis and postsynapsis sequentially. The initial step in presynapsis is the processing of the ends at DSB in 5' to 3' direction to generate 3' single stranded tails. The nucleolytic removal of 5' ends on either side of DSB to generate 3' single strand overhangs is executed by heterotrimeric MRN (Mre11-Rad50-Nbs1) complex (Lamarche *et al.*, 2010). The next sequential step is binding of RPA (Replication protein A) to the single stranded DNA which is subsequently replaced by RAD51 recombinase. There are many paralogues to *RAD51* such as *XRCC3*, *RAD51 B*, *C*, *D*, *XRCC2* (Nowacka-Zawisza *et al.*, 2015). Breast cancer genes *BRCA1* and *BRCA2* are related with *RAD51*. It has been observed that *BRCA2* interacts directly with *RAD51* (Liu and West, 2001). Loss of function of *BRCA1* and *BRCA2* reduces the efficacy of HR (Nagaraju and Scully, 2007). Single stranded DNA with RAD51 is referred to as nucleoprotein filament (figure 2.7). The filament does homology search in the sister chromatid. Once the homologous sequence is

identified Rad51p catalyze the exchange of strands where damaged DNA exploits the DNA duplex and a D-loop is formed (Maher *et al.*, 2011). DNA polymerase η does the gap filling and extension followed by ligase I sealing the nicks. Holiday junction so formed is resolved by cleavage and gaps are ligated. Ku forms ring like structure, where one side of DNA (Kanaar *et al.*, 1998)

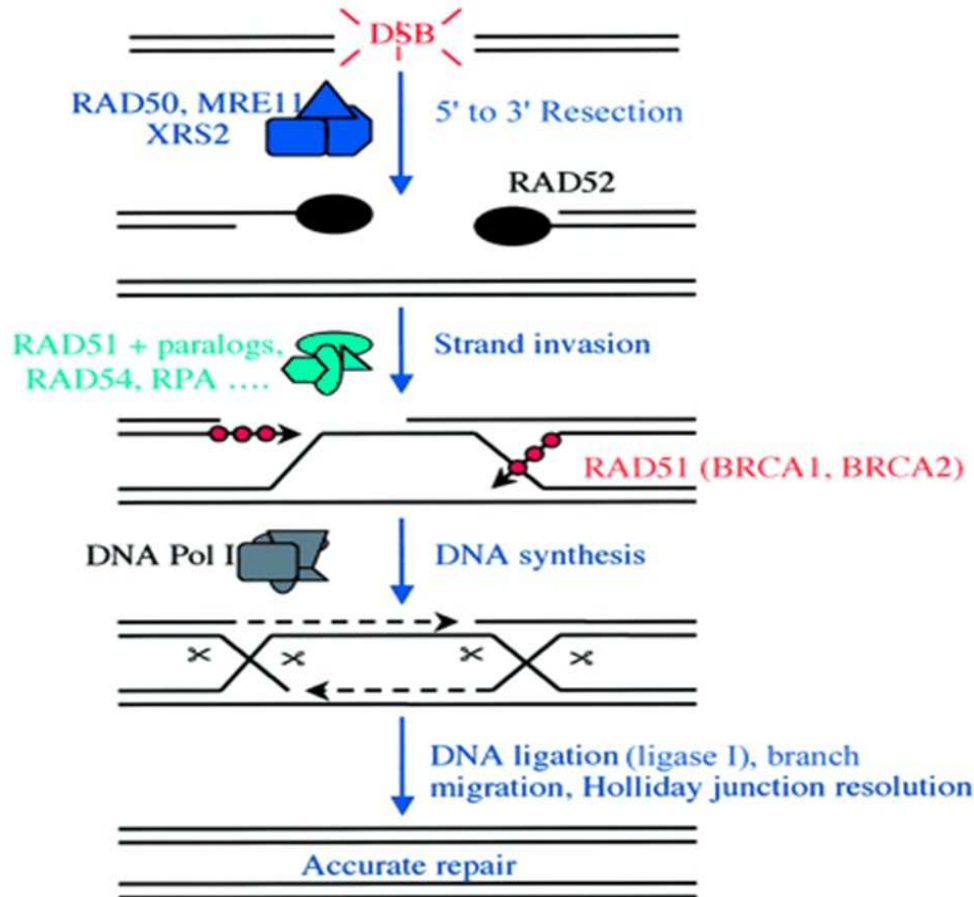


Figure 2.7: Homologous recombination pathway (Source:(Jackson, 2002))

- **Non-homologous end joining-** The pathway is termed as "error-prone" pathway. The canonical pathway starts with Ku protein, a heterodimer of Ku 70 and Ku 80. Ku forms a ring like structure that is disposed on DNA (Rivera-Calzada *et al.*, 2007). The ring forms a cover on one side of the DNA whereas DNA is exposed on the other side allowing for other repair proteins to access the double stranded DNA. Further Ku serves as targeting subunit of DNA-dependent protein kinase catalytic subunit (DNA PKcs) (Chan and Lees-Miller, 1996). DNA PKcs along with Ku forms the DNA-PK holoenzyme. They cause the translocation of Ku allowing the DNA PKcs to come in contact with DNA ends directly (Pawelczak *et al.*, 2005). This signifies that DNA PKcs has affinity for single stranded DNA

as generated due to DSB. DNA PKcs are activated when they come in interaction with single stranded DNA. *XRCC4* and replication factor A2 act as substrate to DNA PKcs. The two DNA molecules generated after DSBs come in close vicinity after DNA PKcs bind to two DNA fragments leading to synapsis (Jovanovic and Dynan, 2006). The DNA ends are modified in case some overhangs are present. Following the sequence, DNA polymerase μ or λ fills the gap. Finally ligation is done by *XRCC4* in conjugation with ligase I as shown in fig 2.8 (Jackson, 2002).

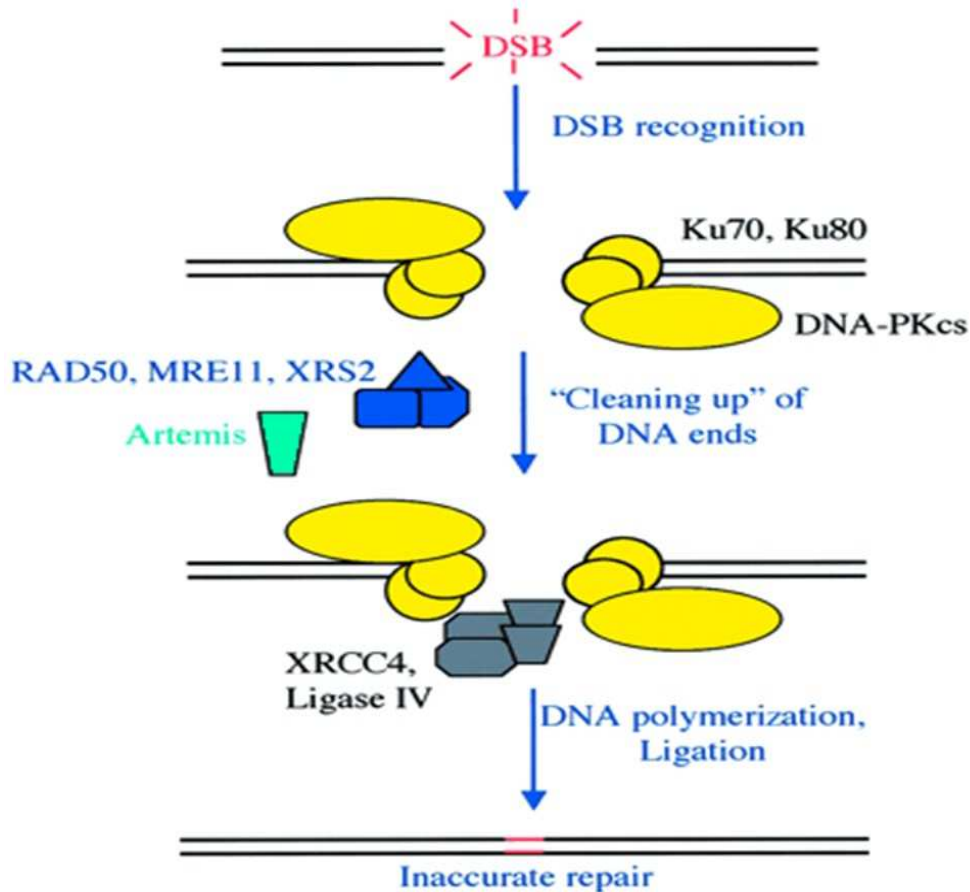


Figure 2.8: Non-homologous end joining pathway (Source:(Jackson, 2002))

There are many important genes in the two sub-pathways:

XRCC3 is an important paralogous gene to *RAD51* with which it shares 20-30% similarity (Nowacka-Zawisza *et al.*, 2015). It is located on chromosome 14q32. It has been studied that *XRCC3* independently associate with DNA breaks and further subsequently helps in the formation of *RAD51* nucleoprotein (Forget *et al.*, 2004). Ku proteins, DNA PKcs are two important proteins of NHEJ subpathway. Ku 70 is also called *XRCC6* whereas catalytic subunit of DNA PKcs is called *XRCC7* respectively. Gene Ku 70 is located on chromosome 22q13 and DNA PKcs is located on chromosome 8q11 respectively (Takiguchi

et al., 1996; Sipley *et al.*, 1995). *XRCC6* is responsible for the maintenance of integrity of genome along with reduction in chromosomal rearrangements. *XRCC4* is located on chromosome 5 at q 14.2 position and codes for a protein of 336 amino acid residue and has 8 exons (Saito *et al.*, 2016). *XRCC4* acts as a connection between Ligase 4 and KU70/KU80. It plays an important role in end joining of blunt DNA ends (Davis and Chen, 2013). Knock-out studies have shown reduced proliferation, chromosomal instability and hypersensitivity to radiations (Gao *et al.*, 2000).

2.5.2 DNA repair and Carcinogenesis

DNA repair forms an integral part in the process of maintenance of genomic integrity. Therefore they are also termed as "gatekeeper" genes which directly influence growth of the tumor may be by promoting growth or death of the cell (Kinzler KW, 1997). Normally cell undergoes its usual growth through cell cycle. However, as and when a damage to DNA increases a certain level, cell cycle check points arrests the cell cycle and stops its progression. DNA repair machinery during this halt performs the repair. Therefore incorrect information is not passed over to the next generation (Houtgraaf *et al.*, 2006). Condition of accumulation of excess of DNA damage leads to shift from DNA repair to apoptosis as indicated in figure 2.9. Inability to undergo apoptosis with unrepaired damages results in enhanced mutations (Williams *et al.*, 1998), which ultimately causes genomic instability (Lengauer *et al.*, 1998). DNA lesions to behave as an active heritable mutations require the premutagens to be replicated or undergo mitosis. When a damaged template is replicated either errors get replicated or blockage of replication fork occurs. The blockades are removed by repair pathways. Lesions in DNA repair coding genes are associated with many diseases. Xeroderma pigmentosum a sunlight induced skin cancer, is the most widely studied hereditary disease. It is caused due to the defects in NER pathway of DNA repair. It results in patients not being able to excise or repair photo-products formed due to UV in skin DNA (Kraemer *et al.*, 2007). Bloom's and Werner's syndrome, Ataxia-telangiectasia *etc.* are other diseases associated. Secondly, a compromised proofreading activity of DNA polymerase δ results in high mutation rate and thereby increased carcinogenesis (Goldsby *et al.*, 2001). Further, DNA repair mostly make use of DNA polymerase β which can bypass certain mutations contributing to carcinogenesis (Srivastava *et al.*, 1999). Adding to the above, observational studies have shown that hereditary non-polyposis colon cancer is associated with defects in mismatch repair genes (Lynch *et al.*, 1993).

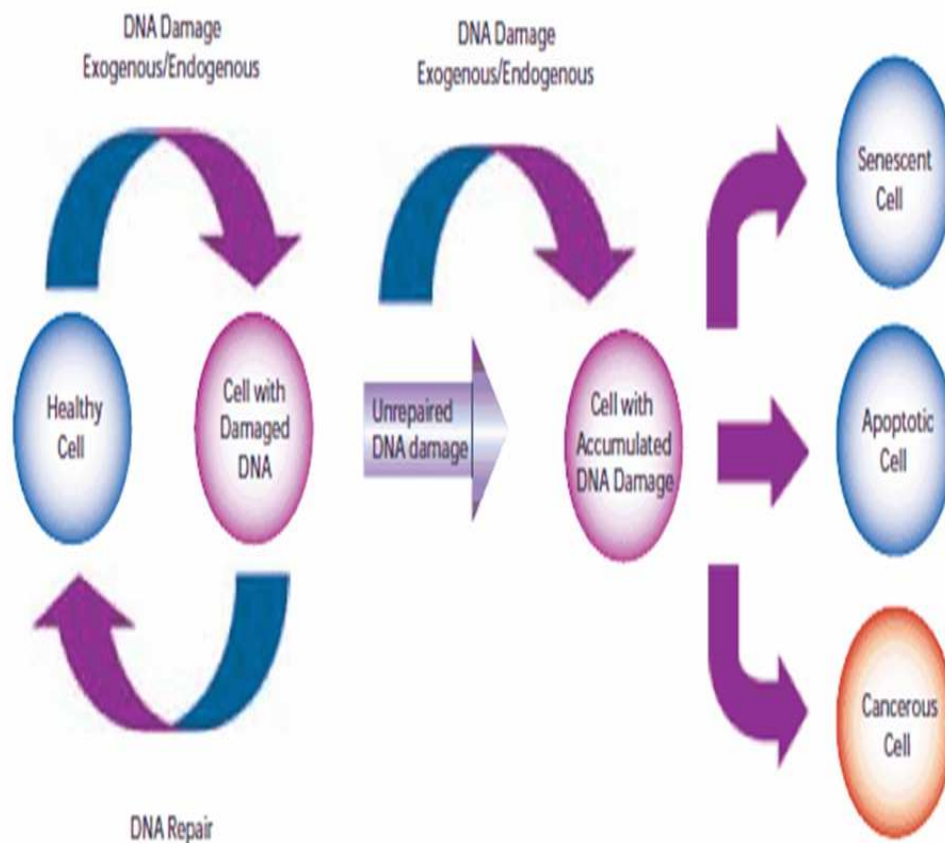


Figure 2.9: Flowchart to cellular response to DNA damage

2.5.3 Genetic variabilities in BER genes: *OGG1*, *MUTYH* and *XRCC1* and their correlation with lung cancer predisposition

Prior data has revealed that there are in total 130 genes involved in DNA repair pathways. SNPs in some play a significant role in carcinogenesis whereas some do not affect the progression to cancer. Various genes are studied in the present work with number of polymorphisms as discussed ahead. Gene *OGG1* has 231 SNPs recorded till date where out of them nearly 25 SNPs have amino acid variants. SNP *Ser³²⁶Cys* (rs1052133) at exon 7 as a result of transversion of C to G at 1245 position, has been studied extensively for its role in cancer. The genotypic and allelic frequency distribution in Taiwan population based study suggested no significant difference in the genotypic distribution between cases and controls followed by significant association of C allele with lung cancer (Liu and Hsia, 2010). Further, study on Taiwan population presented that subjects with double variants in case of *OGG1 Ser³²⁶Cys* acted as a slight risk factor towards lung

cancer. However, the smoking status altered the risk. Never smokers did not show any effect towards lung cancer, moderate smokers showed a mild risk and further the risk amplified for heavy smokers 3.5 times (Chang and Hsiao, 2009). Subjects with double allelic variant genotype in a Chinese study showed a risk towards lung cancer (Qian *et al.*, 2011). Differing with the above results, study on Latinos showed no association of mutant subjects with overall lung cancer susceptibility. On the same lines African-Americans did not show any affect towards lung cancer (Chang, 2008). Similarly, data in Slovak republic showed no association towards lung cancer (Letkova *et al.*, 2013). Combined variant genotype for Thai women was found to be correlated towards lung cancer (Klinchid *et al.*, 2009). However, Chinese population showed nearly no correlation (Liang *et al.*, 2005).

Metanalytic study revealed ethnicity based subgroup analysis for *OGG1* and stated Caucasian population possessed no risk towards lung cancer for mutant genotype in comparison to wild type. Similarly, the combined variant genotype did not show any association with lung cancer (Zhou *et al.*, 2015). Adding to the results, another metanalysis showed mutant showed a mild risk towards lung cancer (Kiyohara *et al.*, 2010). However, both the above met-analytical study also revealed that the subjects with mutant genotype showcased a mild risk in Asian population. Whereas, Asians with combined variant genotype did not show any affect on lung cancer susceptibility (Zhou *et al.*, 2015). Study by Duan *et al.* also did not show any affect on lung cancer susceptibility for both double allelic variant and combined variants. African populace supported the above paradigm (Duan *et al.*, 2012). The most common mutation in *MUTYH* is *MUTYH* association Polyposis (MAP) which is a characteristic of colorectal cancer. However, there is not much data on *MUTYH* in relation to lung cancer. A Japanese study showcased a strong correlation with lung cancer. The minor allele frequency in controls was 0.49 compared to 0.53 in cases (Miyashita *et al.*, 2009). Whereas, Chinese population exhibited a mild association towards lung cancer for the mutant genotype subjects (Qian *et al.*, 2011).

After incision and removal of damage incorporated on DNA. The repair of damage occurs where *XRCC1* plays an important part, as it act as a platform for the activity of DNA polymerase, DNA ligase, and damaged ends of DNA. However there are many polymorphic variants present in *XRCC1* which renders the gene responsible for oncogenic phenotype. The present study includes five important polymorphisms of *XRCC1* i.e *XRCC1 399, 194, 206, 632, 280*.

XRCC1 399 has been immensely evaluated for its role in lung cancer susceptibility in varied population. Chinese subjects for *XRCC1 399* when exhibited a slight correlation (Han *et al.*, 2014) for mutant subjects. Following the trend Mongolian population show-

cased a trend of mild risk (Kang *et al.*, 2015). A Belgium based study however showed a high risk towards lung cancer (De Ruyck *et al.*, 2007). Adding to the study above a significant high risk of 2.5 fold was observed in North Eastern Chinese population (Guo and Li, 2013). Study on Latinos concorded with the above work whereas African America showed almost no affect on Lung cancer (Chang *et al.*, 2008).

Subgroup analysis by ethnicity suggested mutant genotype in Asians were at potent risk of lung cancer (Wang and Cai, 2014). In concordance with the above study Chinese population showed a correlation with lung cancer (Li *et al.*, 2013). However, study by Huang *et al.* showed no association of Asians with lung cancer (Huang *et al.*, 2013). Following the above study, Caucasians also did not show any association with lung cancer both for mutant genotype and combined variant genotype (Wang and Cai, 2014). Another study in Caucasians supported the above study indicating no association of mutant genotype with lung cancer (Huang *et al.*, 2013). African population like the above two populations also did not report any association with lung cancer (Huang *et al.*, 2013). There are many important polymorphisms in *XRCC1* which play a significant role in susceptibility. In the sequence *XRCC1 194* is another important polymorphism studied. Various populations based studies have shown that subjects with mutant genotype for *XRCC1 194* showcased an increased risk towards lung cancer in Chinese population (Han *et al.*, 2014), and stratified Chinese women who daily come in contact with cooking oil fumes (Li *et al.*, 2008). Another study in East Chinese Han population supported the paradigm and gave a high risk towards lung cancer for mutant genotype (Du *et al.*, 2016). In contrary to the above studies there are reports where mutant genotype is responsible for protective (Chang *et al.*, 2008) or no effect towards lung cancer (Guo and Li, 2013). There are studies with very less or no mutant genotype subjects therefore risk is evaluated only for single allelic variant genotype in a Belgium based study (De Ruyck *et al.*, 2007) and European study (Matullo *et al.*, 2005). The stratified population based studies revealed Caucasians showed no risk towards Lung cancer (Wang and Liu, 2014; Yang *et al.*, 2015) for mutant genotype. Whereas mild risk was observed in a metanalysis by Huang *et al.* (Huang *et al.*, 2013). Asian population like prior two studies on Caucasians did not reveal any effect on lung cancer as an odds ratio of 1 was obtained (Wang and Liu, 2014; Huang *et al.*, 2013).

The next very important polymorphism in *XRCC1* is *XRCC1 280* which has a pivotal role in lung susceptibility. Population studies show that overall there was a mild risk observed towards lung cancer however when same metanalysis was stratified for ethnic subgroup it was analyzed Caucasians did not observe any risk, however Asians possessed a risk towards lung cancer (Huang *et al.*, 2013). Another study in the sequence revealed

a significant overall risk towards lung cancer for mutant subjects. Stratification analysis for ethnicity suggested a high risk for Europeans but no risk for Asians (Guo and Yang, 2013). A Chinese study suggested almost no risk towards lung cancer for the same mutant subjects (Han *et al.*, 2014). Amplified risk was observed in case of Chinese women for double allelic variant subjects (Li *et al.*, 2008). However, protective effect was observed in Latinos (Chang, 2008). *XRCC1 206* and *XRCC1 632* are important synonymous polymorphisms recognized, which are less studied in relation to susceptibility. European study revealed no risk for subjects carrying mutant genotype. Subjects with at least one variant allele were at high risk towards lung cancer. Whereas, combined variant genotype for *XRCC1632* did not show any affect on Lung cancer (Yin and Vogel, 2007). Study in Latinos however, showed a mild protective effect towards Lung cancer for mutant genotype.

2.5.4 Genetic variabilities in DSBR genes *XRCC4*, *XRCC6*, *XRCC7* and their correlation with lung cancer predisposition

Polymorphic Double strand break repair pathway genes act lethal to cells as double strand breaks can cause cells to die if not repaired properly. Adding to this there are polymorphisms which are studied for their relevance in many cancers including lung. Gene *XRCC4* is involved in DSBR pathway and acts as a linker between KU70/KU80 and ligase 4. *XRCC4* has shown many polymorphic sites where intron 3 insertion deletion based polymorphism (rs28360071) of 30 bp is included in the present study. Polymorphism based study in Taiwan populace has shown the genotypic and allelic distribution of the polymorphism. There was no significant difference in the genotypic and allelic distribution between cases and controls (Hsu *et al.*, 2009).

Further, *XRCC6* along with *XRCC5* are the initial genes of non-homologous end joining pathway. They recognize the double stranded break and tether other important repair genes required in the process. Polymorphic *XRCC6* in a metanalytic study depicted an overall no risk towards lung cancer for subjects with mutant genotype. However, ethnic subgroup analysis suggested Asians were at risk towards cancer for the same genotype. On the other hand Europeans did not show any risk towards cancer (Jia *et al.*, 2015). Likewise another study did not show any risk towards lung cancer for the mutant genotype. Further, in the study Asians were at risk towards cancer whereas Europeans did not show any risk towards cancer (Jiang *et al.*, 2013). Individual study in Taiwan population depicted no risk towards lung cancer for mutant subjects (Hsia *et al.*, 2012).

Lung cancer in relation to *XRCC7* is not much explored. One popularly known study on Lung cancer is in Taiwan based males and smokers on *XRCC7 G6721T* where mutant subjects showcased a protective effect on Lung cancer. Males and smokers also replicated the protective effect which was highly significant (Hsia *et al.*, 2014). Study on inspection of prevalence of G allele of *XRCC7 G6721T* provided an insight into the frequency of the minor allele in different populations. Caucasians reported it 35-42%. Likewise Hispanic population reported 32% whereas African and African American showed relatively less frequency of 27% (Saadat and Saadat, 2012). Risk analysis for cancer susceptibility suggested overall there was no risk associated with Lung cancer. However analysis for ethnic stratification suggested Asians showed a risk towards cancer whereas, Caucasians did not show any risk (Zhang *et al.*, 2013).

2.5.5 Polymorphic DNA repair genes (OGG1, MUTYH, XRCC1, XRCC4, XRCC6, XRCC7) on Indian population in various cancer

The polymorphic DNA repair gene *OGG1* is observed to have a profound role in Bladder cancer. A 2 fold risk was observed in case of Urothelial bladder cancer in North Indian population. Study on Gallbladder in North Indian population revealed a high risk for the mutant genotype (Srivastava *et al.*, 2009). Above study was followed by another study in Gall bladder on Indian population which showed almost the same level of risk as prior study (Srivastava *et al.*, 2010). The polymorphism when analyzed in head and neck cancer also showed an elevated risk (Kumar *et al.*, 2011). Double allelic variants for the SNP possessed a very strong correlation with Oral cancer in Indian populace (Datkhile *et al.*, 2017). Further, Mittal *et al.* in their study analyzed prostate and bladder cancer together. The outcome of their study stated that the mutant genotype for the polymorphism did not affect Prostate cancer whereas again an amplified risk was observed against bladder cancer (Mittal *et al.*, 2012). Prostate cancer was however, examined in other study on Indian population and no risk was observed (Mandal *et al.*, 2012). When Indian population was analyzed for esophageal cancer and stratified for state based population it was observed for mutant genotype that Kashmiri population did not show any association with cancer however, UP based population were at high risk (Upadhyay *et al.*, 2010). No correlation was observed with colorectal cancer (Sameer *et al.*, 2012).

The gene *XRCC1* is explored immensely in Indian population in relation to various cancers. In this sequence *XRCC1 399* when analyzed in colorectal cancer no significant

association was observed for double allelic variant genotype. However, when study was stratified for gender, males in the study showcased a high risk (Wang *et al.*, 2010). On the contrary a highly protective effect was observed for Gall bladder cancer for mutant genotype (Srivastava *et al.*, 2009). When examined for hepatocellular carcinoma a protective effect was observed similar to bladder cancer (Kiran *et al.*, 2009). Whereas, the same genotypic combination for prostate cancer revealed a high risk (Berhane *et al.*, 2012). Following the trial study by Mittal *et al.* in prostate cancer no risk was observed. The same study also analyzed the risk towards bladder cancer where a high risk was observed (Mittal *et al.*, 2012). Adding to this, introspection into oral cancer revealed a very strong correlation with oral cancer (Ramachandran *et al.*, 2006). Study in breast cancer also revealed a high risk (Datkhile *et al.*, 2017). The previous study also analyzed *XRCC1 194* in breast cancer and no risk was observed. Similarly, study in Gallbladder cancer did not show any risk for the mutant genotype (Datkhile *et al.*, 2017). However, study in prostate cancer depicted a protective effect, on the contrary bladder cancer analysis showed a high risk (Mittal *et al.*, 2012). In parallel to bladder cancer risk evaluation in Oral cancer also depicted a very high risk (Ramachandran *et al.*, 2006).

Studies evaluating *XRCC1 399* and 194 also examined *XRCC1 280* in parallel. Therefore in this sequence it was observed mutant genotype for *XRCC1 280* in hepatocellular carcinoma showed a highly significant risk. Single allelic variant also showed a very high risk (Kiran *et al.*, 2009). On the same lines mutant genotype for oral cancer also showed a high risk (Ramachandran *et al.*, 2006). Whereas study in prostate cancer showed mutant genotype did not show any risk whereas heterozygous genotype showed a high risk. Same study in Bladder cancer revealed similar to prostate cancer no risk for mutant genotype and high risk for single allelic variants (Mittal *et al.*, 2012). Double allelic variants in case of *XRCC1 280* showed highly significant high risk towards breast cancer (Datkhile *et al.*, 2017). Indian population when analyzed for polymorphic DSB gene *XRCC4*, it was observed mutant genotype for *XRCC4* increased the prostate cancer risk (Mandal *et al.*, 2011). When the same genotype was analyzed for urothelial bladder cancer again a high risk was observed for mutant genotype (Mittal and Gangwar, 2012). Likewise on analyzing *XRCC6* in breast cancer in population from Maharashtra no risk was observed for mutant genotype whereas single allelic carriers were highly susceptible towards breast cancer (Patil *et al.*, 2016). A head and neck cancer study showed a high susceptibility towards breast cancer however, single allelic carriers were highly protective (Datkhile *et al.*, 2017). Most renowned study in Indian data related to polymorphic *XRCC7* is on prostate cancer where mutant genotype for the SNP showed 2 fold risk (Mandal *et al.*, 2010). Similarly, *XRCC7* in urothelial bladder cancer showed a very high risk for mutant genotype (Gangwar *et al.*, 2009). On the contrary Breast cancer subjects were

at protective end (Patil *et al.*, 2016).

2.5.6 Genetic variants of DNA repair genes in Lung cancer in Indian population

Polymorphic DNA repair genes are less explored particularly with reference to Lung cancer in Indian population. The literature surveyed till date of the study says polymorphic X-ray Cross Complementing group 1 (XRCC1) was first analyzed by Pachouri *et al.* with a protective effect for polymorphic variant *XRCC1 399* and no correlation with *XRCC1 194* (Pachouri *et al.*, 2007). In the next following study on *XRCC1 399* mutant genotype were at high risk (Sreeja *et al.*, 2008). Later two studies talked about the risk associated with the mutant genotype and both revealed a magnified risk (Saikia *et al.*, 2014; Uppal *et al.*, 2014). There are studies which talk about the distribution of genotypic frequencies of polymorphic XRCC1. Study by Kiran *et al.* showed the genotypic distribution of *XRCC1 399*, *194* and *280*. It was observed the frequencies for 399 and 194 were higher for heterozygous genotype compared to wild and variant genotype. However, in case of *XRCC1 280*, the highest frequency was obtained for wild type genotype (Kiran *et al.*, 2010).

2.6 DNA repair and survival

Lung cancer cases are on continuous rise ever since without any significant improvement in prognosis. However, the maximum 5-year survival rate observed till date is 15-17% (Jemal *et al.*, 2011). Traditionally the treatment options for lung cancer are surgery, platinum based doublet chemotherapy, radiotherapy. Surgery is usually preferred for early stage of lung cancer. For a patient to be a candidate for surgery its necessary to know the type, size, location and grade of tumor along with general health of the patients (Yuan *et al.*, 2015). Studies have shown in this regard, surgery done in patients at early stage have better survival compared to those at later stages. Owing to the limitation of the advanced stage, platinum based chemotherapy is preferred over curative surgery (Spiro and Silvestri, 2005). Though chemotherapy or radiotherapy may be given before or after surgery for effective treatment to occur. The response rate towards platinum based chemotherapy varied between 26-60% (Cui *et al.*, 2012). However, the genetic constituent influences the effectiveness of platinum based doublet chemotherapy.

2.6.1 Association of genetic variants of DNA repair genes with overall survival and prognosis of Lung cancer patients

The expression of variant genotype for the polymorphic variants of DNA repair genes would reduce the efficiency of DNA repair system and thus prolong the effect of chemotherapeutic drugs and ultimately survival. Therefore polymorphic DNA repair genes are observed to influence the survival of population. Henceforth, there is constant effort applied to analyze the prognostic role of DNA repair genes. The present research work summarizes the work done till date on survival in relation to the polymorphic DNA repair genes undertaken in the study.

Cox-regression analysis for BER gene *OGG1* has expressed a strong correlation with double allelic carriers. This hazards ratio increased further when stratified for females (Su *et al.*, 2015). A recent study by Peng *et al.* evaluated Median survival time. It was observed that there was not much difference in the median survival time for mutant and wild type genotype (Peng *et al.*, 2014). Previous study by Su *et al.* also evaluated the influence of *MUTYH* on survival and it was revealed overall there was no association with death rate of subjects with mutant genotype. On the same lines *XRCC1* evaluated in the same study did not show any influence on hazards ratio (Su *et al.*, 2015). Whereas, survival analysis in another study did not express any difference in MST for mutant genotype when compared with the wild type subjects whereas, heterozygous subjects revealed a significant difference in median survival time (Peng *et al.*, 2014). Both *OGG1* and *MUTYH* were analyzed for their involvement in lung cancer prognosis in Japanese population. Univariate analysis for both *OGG1* and *MUTYH* did not show any influence on survival. Multivariate analysis also revealed no risk for death for both the SNPs (Osawa *et al.*, 2012).

One of the initial study on *XRCC1* prognosis came in the year 2004 by Gurubhagavatula *et al.*. It was analyzed the median survival time for mutant genotype was less compared to wild type, indicating a poor survival for mutant genotype. Adding to this, multivariate analysis revealed a very high hazards rate for the same genotype (Gurubhagavatula *et al.*, 2004). Evaluation of genetic variants of *XRCC1* in Japanese populace showed no risk towards death rather, a protective effect was observed for *XRCC1 399* (Osawa *et al.*, 2012). Further many studies supported the above study where a protective effect was observed. Study by Giachino *et al.* (Giachino *et al.*, 2007), Ke *et al.* (Ke *et al.*, 2012), Zhang *et al.* (Zhang and Ma, 2014), Jin *et al.* (Jin *et al.*, 2014) showed a very less hazards rate which depicted a better survival for mutant genotype of *XRCC1 399*. While there are studies which report no correlation with survival. Study on Chinese population

(Yao *et al.*, 2009), (Yuan *et al.*, 2010), (Zhao *et al.*, 2013) revealed a lack of influence on prognosis of lung cancer for double allelic variants.

XRCC1 194 similar to *399* showcased a contradictory results where some studies exhibited a high hazard rate and some showed no correlation or high death rate. Double allelic variants for *XRCC1 194* exhibited a significant protective effect. The adjusted hazards rate for the same was very small signifying very small death rate (Zhao and Chen, 2015). Similarly, a Chinese study expressed a very small hazard rate (Ke *et al.*, 2012). Whereas, no association was observed for the mutant subjects in other Chinese studies (Yuan *et al.*, 2010; Zhang and Ma, 2014; Zhao *et al.*, 2013). The next important SNP studied in the present study is *XRCC1 280*. Analysis of *XRCC1 280* gave a trend of lesser hazard rate comparatively (Zhao and Chen, 2015). Whereas another study exhibited no influence on death rate (Jin *et al.*, 2014).

Polymorphic X-ray Cross Complementing group 3 (*XRCC3*) is well known for its adduct formation. It has also been explored for its role in prognosis related to lung cancer in varied populace. Genotypic variant of *XRCC3* is evaluated for survival in Poland based populace, Chinese population where a very high Hazard rate for mutant and combined variant subjects was observed (Butkiewicz *et al.*, 2011; Ren *et al.*, 2012). Likewise, Japanese population also revealed a strong correlation with survival. There are reports which exhibit no effect towards death rate (Ke *et al.*, 2012). On the same lines, a Chinese study showed no influence on death rate. *XRCC3* is observed to have important role in clinical outcomes of platinum based chemotherapy used for treatment of lung cancer than survival (Qiu *et al.*, 2013). Further, there are very rare reports available on polymorphic DSB repair genes such as *XRCC6* and *XRCC7* in relation to survival.

Polymorphic DNA repair genes influence survival, as variant genotype may be responsible for better activity of chemotherapeutic drugs hence better survival. There are very rare few reports on implication of variants of DNA repair genes in predicting the prognosis and survival of lung cancer patients in Indian population. However, there are two studies for polymorphic *XRCC1* in context of survival. Study by Syamala *et al.* talked about survival in Breast cancer. Familial Breast cancer for combined variant genotype in case of variant *XRCC1 399* showcased no risk. Whereas, same genotypic combination for Sporadic breast cancer exhibited a very small death rate (Syamala *et al.*, 2009). The other study on survival was in lung cancer where a strong correlation was observed with survival (Sreeja *et al.*, 2008).

Chapter 3

Gaps in study and objectives

3.1 Study population

There have been very few reports on known gene of DNA repair pathway *i.e.* *XRCC1* in North Indian population. To the best of our knowledge there have been no reports in rest of the BER and DSB genes in lung cancer in North Indian population. Also none has reported the combined effect of the repair genes polymorphism towards lung cancer risk. In addition, there have been no studies conducted in Indian population regarding the role of the SNPs in the DNA repair genes and its association with clinical outcomes of lung cancer patients undergoing platinum based chemotherapy. It has been reported if chemotherapy cannot control the tumour growth, the detrimental effects of chemotherapy may reduce quality of life of the patients. As many DNA repair genes carry polymorphisms with the potential to modulate gene function altering DNA repair capacity and impairing the DNA repair mechanisms which are involved in removing DNA adducts, therefore identification of polymorphism of DNA repair genes and understanding their correlation to survival in lung cancer patients and susceptibility to lung cancer would yield insights how the gene polymorphism adversely or favourably influence the chemotherapy outcomes in the north Indian population. Furthermore these SNPs could provide targets for the future development of personalised therapeutics.

3.2 Objectives of the study

To study the single nucleotide polymorphic variants of BER and DSB genes by Polymerase chain Reaction-Restriction Fragment Length polymorphism (PCR-RFLP).

- To study the single nucleotide polymorphic variants present in the Base Excision repair (BER) system.
- To study the genetic polymorphism present in the Double Strand Break (DSB) repair genes.

- To measure susceptibility of different DNA repair genes polymorphism with lung cancer development.
- To test for association between gene polymorphism and clinical outcomes in patients receiving platinum based doublet chemotherapy.

Chapter 4

Material and Methods

4.1 Study population

Lung cancer cases and controls enrolled in the present thesis study, were registered at the Department of Pulmonary Medicine, PGIMER Chandigarh. This study has been accepted by the Institute ethics committee of PGIMER. Ethical clearance letter is attached in the Appendix-I. A written consent was obtained from each enrolled subject and their representatives. Informed written consent was obtained from all participants as attached in Appendix-II.

The enlisted patients were histopathologically diagnosed with NSCLC or SCLC subtype. The study included the newly diagnosed lung cancer patients with no age, gender, stage, histology and TNM restrictions. Patients under observation or with a prior history of cancer were excluded from the study. Controls included in the study are people who have come for normal health checkup and do not have any medical record in relation to cancer at the time of blood collection. Healthy controls were selected in a manner so as to pairwise match the age (10 years), gender and smoking habits with the cases. This was done in order to avoid any sampling bias.

Patients meeting all the following requirements were eligible for enrollment:

(1) Diagnosis of NSCLC & SCLC either confirmed by cytology or histology (2) Stage IIIB or IV disease, (3) Untreated and intent to treat with definitive chemotherapy, (4) An Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 03, (5) At least one bi-dimensionally measurable lesion, according to the RECIST criteria, (6) adequate organ function, defined as absolute neutrophil count $>1500/\mu\text{L}$, platelet count $>100,000/\mu\text{l}$, and levels of creatinine.

Criteria for patient withdrawal from the study were: unacceptable toxicity, non-cancer-related illnesses, which might preclude continuation of treatment, or on patient request. All patients were clinically staged based on physical examination, plain chest radiograph and contrast enhanced computed tomography (CT) scan of thorax and upper abdomen (including liver and adrenal glands). PET-CT ultrasonic examination of abdomen and

pelvis CT/MRI scan of brain, isotope bone scanning or other relevant investigations were undertaken if clinically indicated. Age, gender, smoking habits were recorded. Around 5ml blood is collected from every study subject.

Smoking status divides the respondents in three categories: Current smoker, Former smoker and Never smoker. Current smokers were those who have smoked 100 cigarettes in their lifetime and are currently smoking. Former smokers who have smoked at least 100 cigarettes in their lifetime and who at the time of survey, did not smoke at all. Respondents who reported never having smoked 100 cigarettes were defined as Never smokers. We have divided our study subjects into two groups smokers (current smokers) and non-smokers. Pack-years for smokers were obtained using this equation: [(cigarettes or beedis (type of Indian cigarette) per day/20) X number of years smoked].

4.2 Chemotherapeutic regimens

Doublet platinum based chemotherapy was administered to all lung cancer subjects enrolled in the study. Study subject were given chemotherapeutic drugs cisplatin or carboplatin in combination with either of them docetaxel, irinotecan or pemetrexed. The combination chemotherapy consisted of docetaxel 75mg/m², 500mg/m² pemetrexed or 100mg/m² irinotecan (administered as a 1-h infusion ,followed by cisplatin 65mg/m² administered over 3-hrs as an intravenous infusion). A total of 162 patients were administered the chemotherapeutic treatment. All chemotherapeutic drugs were administered intravenously, and treatment cycles were repeated every 34 weeks. All patients received standard folate and Vitamin B12 supplementation. In presence of unacceptable toxicity or clinico-radiological suggestion of disease progression, tumor response assessment were performed prior to completion of four cycles and chemotherapy was stopped, if indicated. Patients with an objective response to treatment received two additional cycles (maximum of six cycles).

4.3 Follow up and response assessment

All patients were followed up on telephonic conversation every two months until death or the end of the study period of 3 years. Survival time was considered from the date of diagnosis till the date of last follow up or death due to other reasons. Our primary end point was overall survival and the second end point was relationship between genetic polymorphism and treatment responses. Tumor response assessment was evaluated using

Response Evaluation Criteria in Solid Tumors (RECIST) criteria. By using RECIST criteria responses were broadly classified into four categories: complete response (CR), partial response (PR), stable disease (SD), and progression disease (PD). For analysis, we grouped CR and PR as responders and SD and PD as non-responders.

4.4 Genomic DNA isolation

The genomic DNA is isolated using standard Protein K digestion, phenol/chloroform extraction and ethanol precipitation method from blood samples (Field *et al.* 1999) using the protocol given below.

- Equal amount of blood and washing Buffer (5 ml) (1.6 ml 1M Sucrose, 0.5 ml Triton X-100, 0.25ml $MgCl_2$, 0.5 ml 100mM Tris HCl and 0.26 ml of water) are added and mixed thoroughly and centrifuged at 3500 rpm for 5 minutes.
- The supernatant is discarded and 5ml of washing buffer is added.
- The pellet is re-suspended followed by centrifugation. The step is repeated thrice.
- The pellet is dissolved in 1ml of lysis buffer (1 M Tris HCl 2ml, 10% SDS 0.5ml, 0.5 M EDTA 0.6ml, 5M NaCl 0.15ml, 10mg/ml Proteinase-K 0.05ml and water 1.7ml) and incubated at 55°C overnight).
- Next step is addition of phenol:chloroform:isoamyl alcohol (PCI) in the ratio of 25:24:1 in equal volume i.e. 5 ml.
- Centrifugation is done at 8000 rpm for 10 minutes.
- The above two steps are repeated.
- The supernatant is aliquoted and add equal volume of chloroform:isoamyl alcohol (24:1) solution is added.
- Centrifuge it at 6500 rpm for 5 minutes and the upper layer is taken.
- To the aqueous upper layer, equal volume of chilled Isopropanol or 2.5 times volume of absolute Ethanol is added and mixed slowly.
- Centrifugation at 12,000 rpm for 10 min at 4°C
- The supernatant was discarded and the pellet of DNA was washed with chilled 70% Ethanol twice at 10,000 rpm for 5 minutes
- Decant ethanol and air dry the pellet

- The pellet is dissolved in 50-150 μ l EDTA buffer depending on the size of pellet.

4.5 Qualitative and Quantitative estimation of DNA

DNA after isolation is quantified using Nanodrop Spectrophotometer (Thermoscientific). The nanodrop is such designed that it can measure the nucleic acid concentrations in sample volume of 1 microliter. This is advanced spectrophotometric technique as there is no need for cuvettes and capillaries. Sample is directly placed on detection surface and surface tension creates a column between the ends of optical fibres. The absorbance of DNA is noted at two wavelengths A260 nm and A280 nm. To calculate the purity of DNA, ratio of absorbance at 260 nm and 280 nm is used. If A260 nm/ A280 ratio is approximately 1.8 then the DNA sample is pure and free from contaminations of RNA and proteins. If A260 nm/ A280 ratio is approximately 2.0 or >1.8 it shows RNA contamination in the sample, whereas a ratio less than 1.8 shows protein contamination in the sample. The concentration of pure double stranded DNA at 1 OD 50g/ml. The DNA concentration in a solution was determined by using this formula given below: Concentration (g/ml)= A260 X 50 g/ml X Dilution factor.

The isolated DNA is checked or its quality by looking for intact bands of DNA on agarose gel. 0.8% agarose gel was run with 0.5 μ g/ml ethidium bromide immersed in 0.5X TBE. The sample was mixed with 6X-loading dye and diluted with sterile water before loading. The gel was visualized under UV-transilluminator.

4.6 Genotyping of polymorphic sites of BER and DSBRR genes

4.6.1 Genotyping of BER genes

The SNPs of the genes in the study are genotyped using PCR-RFLP (Polymerase chain reaction Fragment Length Polymorphism) method. The genotyping of five variants of *XRCC1* along with that of *MUTYH* was carried out by PCR-RFLP. The primer sequence for *OGG1* were as detailed by Ding *et.al* (Ding *et al.*, 2015). The primer sequence of *MUTYH* were similar as that discussed before by Synowiec *et.al* (Synowiec *et al.*, 2012). The primer sequence of the five polymorphic variants of *XRCC1* was similar to that as described previously (Lunn *et al.*, 1999; Yin and Vogel, 2007; Zhang and Wan, 2005). The

primer sequence, amplified product size, annealing temperature are mentioned in table 4.1. Volume of PCR mixture reaction was 25ml and it consisted of 1x PCR buffer, 1.5mM Mgcl₂ with 0.5 μ M of both forward and reverse primer, 200 μ M of each dNTP's, 100 μ g/ml bovine serum albumin (BSA) and 1U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific) and approximately 200ng DNA.

Table 4.1: List of Primers and restriction enzymes used in the study

Gene, SNP ID	Exons, (Nucleotide and Amino acid substitution)	Primer sequence	Annealing temperature	Product size(bp)	Restriction enzyme	Restriction pattern
XRCC1 194 rs 1799782	Exon 6 C > T(Arginine-Tryptophan)	F5-CCC TTT GGC TTG AGT TTT GT-3R5-TCA GAC CCA GGA ATC TGA GC-3	52 ⁰	375	<i>PvuII</i>	375bp=C 254+121=T
XRCC1 206 rs 915927	Exon 7 A > G(Proline-Proline)	F5-GTC CCA TAG ATA GGA GTA AAA-3R5-CCC TAG GAC ACA GGA GCA CA-3	60 ⁰	432	<i>MspI</i>	432bp=A 344+88=G
XRCC1 280 rs 25489	Exon 9 G > A (Arginine-Histidine)	F5-TGG GGC CTG GAT TGC TGG GTC TG-3R5-CAG CAC CAC TAC CAC ACC TGA AGG-3	60 ⁰	280	<i>RsaI</i>	280=G 280+140=A
XRCC1 399 rs 25487	Exon 10 G > A(Arginine-glutamine)	F5-TTG TGC TTT CTC TGT GTC CA-3R5-TCC TCC AGC CTT TTC TGA TA-3	62 ⁰	615	<i>MspI</i>	615bp=A 375+240bp=G
XRCC1 632 rs 3547	Exon 17 G > A(Glutamine-Glutamine)	F5-GCC AAG CAG AAG AGA CAA A-3R5-GAG TGG CTG GGG AGT AGG A -3	62 ⁰	556	<i>HaeIII</i>	256+192+108 bp = G 256+300=A
OGG1 326 rs 1052133	Exon 7 C > G (Serine-Cysteine)	F5-CTG TTC AGT GCC GAC CTG CGC CGA-3R5-ATC TTG TTG TGC AAA CTG AC-3	60 ⁰	247	<i>MBOI</i>	224+23bp=C 247=G
MUTYH 324 rs 3219489	Exon 12 G > C (Glutamin-Histidine)	F1 5-CCT GTC GGG CAG TCC TGA CG-3 R1 5-CGC TGA AGC TGC TCT GAG GGC-3 F2 5-CCC AGC TCC CAA CAC TGG ACA C-3 R2 5-GAG GCA GGC ACA GGT GGC AC-3	65 ⁰	562	-	241bp=G, 362bp=C

4.6.2 Restriction Fragment Length Polymorphism

After the PCR amplification, the amplicon is treated with respective restriction enzyme. 10 μ l of amplified product was digested with 5units of restriction enzyme. A 20 μ l reaction was set in the presence of 1X respective enzyme buffer. The reaction was kept for incubation at the desired temperature which is 37°C for 10-12 hours.

Resolution of digested product was done on 2.5% agarose gel dissolved in 0.5X TBE with 0.5 μ g ethidium bromide. 6X loading dye containing sucrose, bromophenol blue and xylene cyanol, was mixed with digested product before loading. UV-transilluminator is used to visualize the resolved gel. Scoring of gels was done by two independent persons to avoid any biasness.

The restriction pattern was different for every polymorphic variant and accordingly scoring of the gels was performed. 15% of samples were randomly genotyped twice to confirm the reproducibility of the results.

4.6.3 Polyacrylamide gel electrophoresis

PAGE (Polyacrylamide gels) is used to separate nucleic acids and proteins in specific buffer under electric field. PAGE gel consists of acrylamide and bis-acrylamide. Acrylamide molecules joined head to tail with each other and bis-acrylamide cross-links the two monomeric acrylamide molecules. The two components are mixed in a fixed proportion which helps determine the pore size of the gel. The polymerization in PAGE starts with APS (ammonium persulphate) and catalyzed by TEMED (N, N, N, N-tetramethylethylenediamine). TEMED accelerates the generation of free radicals from APS, which ultimately is responsible for the polymerization of PAGE gel. PAGE gives a very high resolution of DNA molecules and a separation of fragments differing by 1bp is also possible. After vertical gel is run silver staining is done to make the bands on gel visible. Usually a 30% solution of gel is made making use of 29:1 ratio of acrylamide and bis-acrylamide. It further requires 5X TBE, 10% APS, TEMED and 6X loading dye. The study requires 6% gel for resolution of DNA fragments.

4.6.4 Silver staining

Silver staining is a sensitive technique for staining of nucleic acids or proteins in PAGE gels. The nucleic acid and protein bands are visible in the technique after the polyacrylamide gel is impregnated in soluble silver ion (Ag¹) solution. The macromolecule present

in the gel help to reduce the soluble silver (Ag^1) to insoluble metallic silver (Ag^0) which deposits on the gel making the bands visible. The deposition of metallic silver is an auto-catalytic process and promotes further deposition of silver metal increasing the sensitivity of the process. Steps involved in the process make use of solution listed below:

- Fixative Solution: Water:Methanol:Glacial acetic acid (50:40:1)
- Staining solution: 0.1gm of AgNO_3 in 100 ml of distilled water. $150\mu\text{l}$ of 37% is also added to it.
- Developing solution: 3gm of Sodium Carbonate in 100 ml distilled water and $150\mu\text{l}$ of 37% formaldehyde and $20\mu\text{l}$ of 10mg/ml sodium thiosulfate.

Methanol and Glacial acetic acid used for the purpose of fixing the gel, as this make the macromolecules insoluble and preventing it from diffusing out of the gel. The major action of reduction of silver to silver metal is carried out by formaldehyde at high pH which is maintained because of sodium carbonate used in developing solution. The gel is fixed in fixative solution for 30 minutes minimum followed by staining for next 30 minutes. Final step is developing where gel is kept at shaking after pouring the developing solution. The bands developed and observed. In between every step gel was washed thrice with deionized water.

4.7 Genotyping of DSB_R genes

PCR-page was performed for genotyping of *XRCC3*. PAGE done for the visualizing purpose was similar to as detailed in genotyping for *OGG1* variant. The next SNP of DSB_R pathway was *XRCC4* for eich the primer sequence was similar to (Yin and Liao, 2012). Here insertion or deletion was observed by presence or absence of band on a agarose gel. The presence of band of 139bp indicated insertion and presence of 109bp suggested deletion. The amplified product is run on 3% agarose gel and visualized under UV-illuminator. *XRCC6* and *XRCC7* genotyped by PCR-RFLP. The primer sequence, product size restriction enzyme and digestion pattern is enlisted in the table 4.2. The primer sequence used for amplification was similar to that described before (Jahantigh *et al.*, 2015).

Table 4.2: List of Primers and restriction enzymes used in the study

Gene, SNP ID	Exons, (Nucleotide and Amino acid substitution)	Primer sequence	Annealing temperature	Product size(bp)	Restriction enzyme	Restriction pattern
XRCC3 241 rs 861539	Exon 7 C > T (Threonine-Methionine)	F5'-GCT GTC TCG GGG CAT GGC TC-3' R5'-ACG AGC TCA GGG GTG CAA CC-3'	67 ⁰	208	<i>NlaIII</i>	208=C, 120+88=T
XRCC4 rs 28360071	Intron 3 Insertion/Deletion	F5'-TCC TGT TAC CAT TTC AGT GTT AT-3' R5'-CAC CTG TGT TCA ATT CCA GCT T-3'	50 ⁰	-	-	Insertion =139, Deletion=109
XRCC6 61 rs2267437	promoter C>G	F5'-TCT CCA CTC GGC TTT TCT TCC A-3' R5'-TCT CCC TCC GCT TCG CAC TC-3'	67 ⁰	320	<i>BANI</i>	262+58=C, 182+80+58=G
XRCC7 6721 rs 7003908	Intron 48 G>T ()	F5'-CGG CTG CCA ACG TTC TTT CC-3' R5'-TGC CCT TAG TGG TTC CCT GG-3'	62 ⁰	368	<i>PvuII</i>	368=G, 274+94=T

4.8 Statistical analysis

Demographic analysis: Chi square test was used to compare the cases and controls for the selected demographic characteristic such as gender and smoking status. Whereas t-test was used to compare groups with regard to continuous variables such as age and smoking.

Genotypic and allelic Frequency analysis: The HardyWeinberg equilibrium theory ($p^2+2pq+q^2=1$; where p is the frequency of the wild-type allele and q is the frequency of the variant allele) was used both in cases and controls to calculate the genotype frequencies of all SNPs using χ^2 test. Minor allele frequency was also calculated. Pearson's χ^2 test was used to determine whether there was any significant difference in allelic and genotypic frequencies between cases and controls. The minor allele frequency was also calculated for each polymorphic site.

Association analysis: Odds ratio (ORs) and 95 % CI were calculated to assess the risk for lung cancer and the polymorphisms. Adjusted Odds Ratio (ORs) were calculated using logistic regression for subsequent analysis of potential risk factors such as age, smoking, sex. All the variables used in the study were dichotomous except the continuous variables such as age. In the logistic regression analysis, the occurrence of lung cancer was taken as the dependent variable and genotype was selected as the independent variable which can predict the susceptibility for lung cancer. In addition to overall association analysis, stratified analysis was performed to estimate risk for subgroups by smoking status and tumor histology. All the above statistical analyses were performed with Medcalc version 15.11.4 (Medcalc Software, Ostend, Belgium) and SPSS Version 20.0.

Haplotype analysis: Haplotype frequency estimation and linkage dis-equilibrium between the groups was analyzed using SHEsis software (<http://202.120.31.177/myAnalysis.php>) (Yong and Lin, 2005). The polymorphic sites which share a common chromosomal loci and belong to the same gene were evaluated for LD in order to obtain D and r^2 values. Only those haplotype blocks were evaluated where the case and control frequency was more than 0.03, rest of the haplotypes were not considered for finding their correlation with lung cancer predisposition. The difference in the frequencies was calculated using global chi square test. The association was estimated by finding out the odds ratio for each haplotype.

False discovery rate analysis: Multiple comparison corrections were carried out for genotypic combinations (SNP- SNP) using FDR (False discovery rate) method by on-line software based on Benjamini and Hochberg approach (<http://sdmproject.com/util->

ities/?show=FDR). The adjusted p values are mentioned and FDR adjusted $p < 0.1$ was considered as significant (Benjamini and Hochberg, 1995).

MDR: We have applied the multifactor dimensionality reduction (MDR) method to identify interaction models. MDR is a non-parametric genetic model method for overcoming some of the limitations of logistic regression for the detection and characterization of SNP interactions. MDR make use of cross validation consistency (CVC) and permutation testing for revealing the disease status. CVC requires the division of data into training and testing set. The analysis is performed ten times generating the interaction models. Using MDR multi-locus genotypes are pooled into two groups, high and low risk, thus reducing the genotype predictors from n dimensions to one. Among the different genotype models generated only those genotypic combinations, having the highest cross-validation consistency (CVC) and minimum prediction error were taken as the best interaction model. MDR tests were performed using the version 0.5.1 of the open-source MDR software package that is available online (<http://www.epistasis.org>) (Ritchie and Motsinger, 2005).

CART analysis: Lastly, we conducted a classification and regression tree (CART) analysis to detect and characterize the high-order interactions employing CART software (6.0, Salford Systems, San Diego, California). CART is a tree based model that is created by binary recursive partitioning method and produces a decision tree to identify sub-groups at higher risk, which are found to be less visible when using logistic regression methods. The analysis is conducted in such a manner where the most significant predictor is used to split the sample into subgroups and continues until the differences are not significant. Finally, the results are obtained as classification or decision, trees having a node and sub-nodes. The risk of all genotypes sets was estimated by considering the node with low case rate was as the reference to calculate the ORs and 95% CIs (Gupta *et al.*, 2016).

Overall survival and Hazard rate analysis: Association between overall survival and genetic polymorphism was evaluated using Kaplan Meier and Log rank test. Comparison of survival curves was accessed using log-rank test. Cox proportional hazard model was used to calculate the hazard rate and assess the effect of genetic polymorphism on overall survival after adjusting for other covariates. Hazard ratios (HRs) and 95% CIs were estimated using Cox proportional hazards models. In this model, the response (dependent) variable is the hazard. The hazard is the probability of dying that patients have survived upto a given point in time or the risk for death at that moment. Logistic regression model was used to evaluate relation between response and genetic polymorphisms, adjusted for age, gender, smoking, stage, histology, and performance status and chemotherapy regimen. All tests were two-sided and statistical significance was set at

$p < 0.05$. All the statistical analysis was performed with Medcalc version 16.8.1 (Medcalc Software, Ostend, Belgium).

STREE analysis: Further, high order gene interactions were taken into account by using the recursive partitioning approach. This approach produces the initial split taking into consideration the primary factor affecting the OS of patients. The software used for making the decision tree is STREE program (<http://c2s2.yale.edu/software/stree>). Log rank method was employed as the splitting criteria. The recursive method keeps on producing the offspring nodes till statistically insignificant split occurs. The resulting tree was binary, and each terminal node represented a subgroup of patients carrying a particular genotypic combination and hence shows a varied survival time and prognosis. HRs and their 95% CIs for each terminal node were calculated using the multivariate Cox proportional hazard model adjusted for sex, age, histology, KPS, ECOG, smoking and stage of tumor.

Chapter 5

Results

5.1 Genomic DNA Isolation

Genomic DNA was estimated both qualitatively and quantitatively. Qualitative estimation of genomic DNA was carried out using 0.8% agarose gel. Bands were obtained and checked for integrity. Figure 5.1 represents agarose gels for genomic DNA for both cases and controls.

5.2 PCR amplification

Amplification of gene was carried out by PCR. Each amplified SNP was confirmed by running the PCR product on 1.7% agarose gel. The product size of the amplified product was confirmed by running a DNA marker of either 100 bp or 50 bp along with PCR product.

5.3 Genotyping of polymorphic variants by RFLP

5.3.1 Genotyping for BER genes

Agarose gel pictures elucidating the amplified DNA product and restriction pattern for polymorphic variants of BER genes are given below in figures 5.2 and 5.3. Polymorphic *OGG1 Ser³²⁶Cys* was analyzed which is located in the exonic region of the gene. This polymorphism is responsible for amino acid change from Serine to Cysteine in exon 7. The amplified PCR product is of 247 bp. The restriction enzyme used for restriction digestion is *MboI*, the genotype gives a band of CC (224, 23 bp) for wild type, CG (224, 23, 247 bp) for heterozygous, GG (247 bp) for mutant genotype as exhibited in figure 5.2a. *MUTYH Gln³²⁴His* is other polymorphic variant of BER pathway analyzed in the present study. The PCR-CTTP method was used to amplify the product. A common

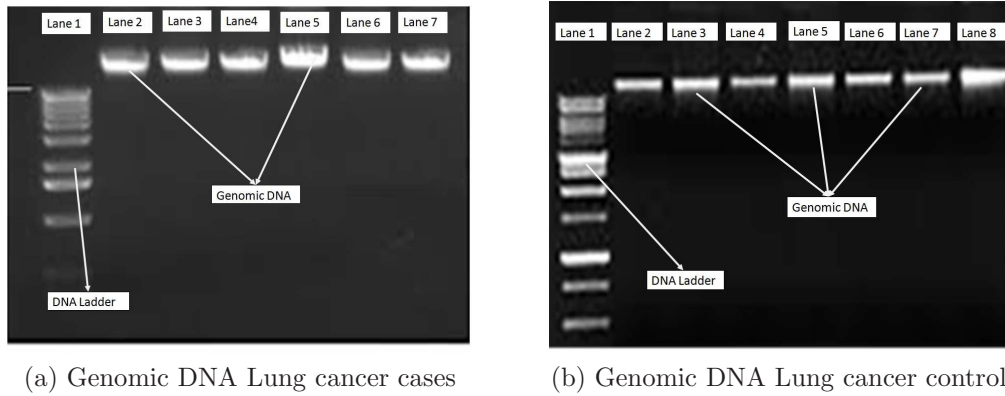


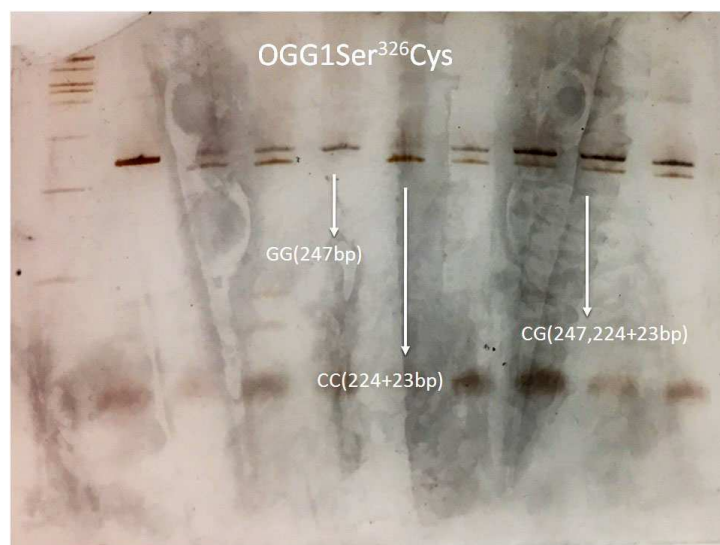
Figure 5.1: Representative example of agarose gel electrophoresis showing genomic DNA bands. (a) Genomic DNA of lung cancer cases. Lane1 represents the DNA ladder, Lane 2-7 represents the genomic DNA in lung cancer cases (b) Genomic DNA of lung cancer controls. Lane1 represents the DNA ladder, Lane 2-8 represents the genomic DNA in lung cancer controls

PCR product of 562bp was obtained. When wild allele is expressed a band of 241 bp was observed along with 562 bp and when variant allele is expressed a band of 362 bp is seen whereas heterozygous were depicted with three bands of 562, 362 and 241bp as depicted in figure 5.2b.

The five polymorphic sites of *XRCC1* were also analyzed further in the study. Single nucleotide polymorphism *XRCC1 Arg³⁹⁹Gln* is located in exon 6 and resulted in amino acid substitution. The polymerase chain reaction produced a product of 615 bp and *MspI* was used for restriction digestion. The wild genotype *GG* was cut into two products of 375, 240 bp while the mutant genotype *AA* remained uncut and was of 615 bp. The heterozygous genotype had three bands of 615 bp, 375 bp, and 240 bp as represented in figure 5.2c . *XRCC1 Arg¹⁹⁴Trp* was also analyzed as indicated in figure 5.2d. This SNP is located in exon 6 and an amplicon of 375 bp is produced. *PvuII* was used for restriction digestion of the amplified product and gave this pattern: *CC* (375 bp) for wild type genotype, *TT* (254, 121 bp) for mutant genotype and *CT* (375, 274, 121 bp) for heterozygous genotype. Next SNP is *XRCC1 Pro²⁰⁶Pro* is present in exon 7 and it is a synonymous polymorphism that results in A-G substitution. The amplified product of *XRCC1 Pro²⁰⁶Pro* is of 432 bp and is cut by *MspI*. The digestion pattern obtained exhibited wild type genotype of 432 bp and mutant subjects of 344 bp and 88 bp as indicated in figure 5.3a.

We have also analyzed the *XRCC1 Gln⁶³²Gln* which is in exon 17 of *XRCC1* gene and results in G-A substitution leading to *Gln⁶³²Gln*. The amplicon is of 556 bp and digestion by *HaeII* gave a wild type with three bands of 256, 192, 108 bp respectively whereas,

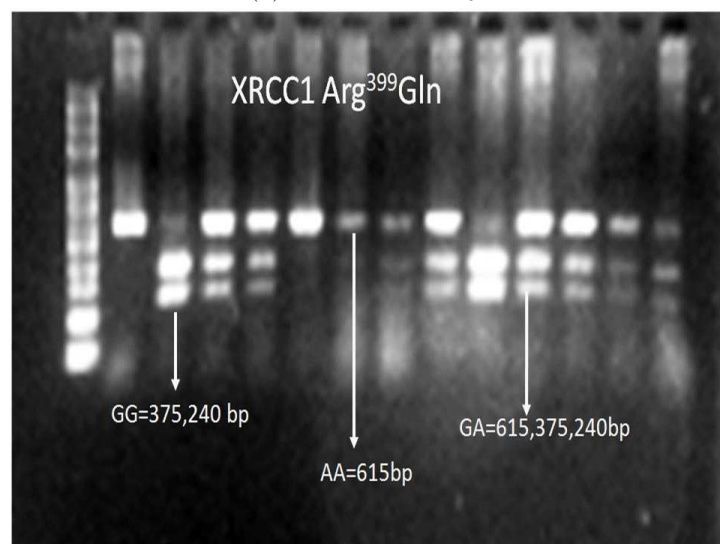
mutant subjects were of two products 256 bp and 300 bp as shown in figure 5.3b. The last SNP studied of *XRCC1* is *XRCC1 Gln²⁸⁰His*. The restriction pattern of amplified pcr product is wild type (280 bp), mutant type (280, 140 bp) as exhibited in figure 5.3c. *RsaI* was used for the digestion of amplified product of *XRCC1 Gln²⁸⁰His*.



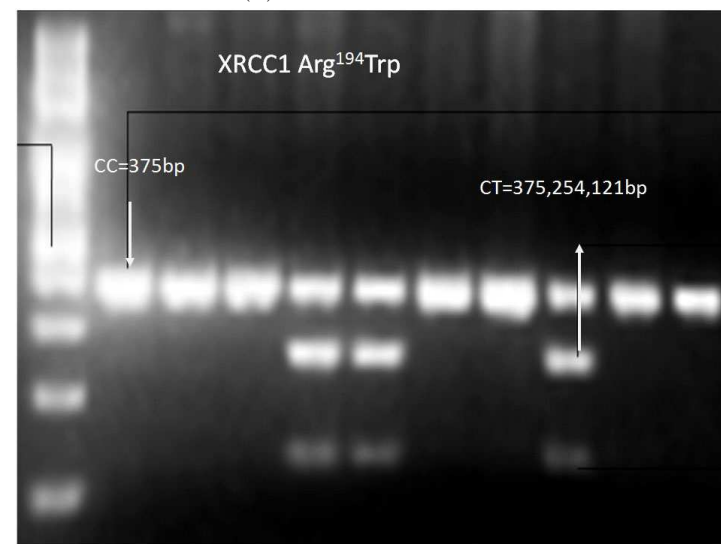
(a) *OGG1 Ser³²⁶Cys*



(b) *MUTYH Gln³²⁴His*



(c) *XRCC1 399*



(d) *XRCC1 194*

Figure 5.2: Representative examples of agarose gel electrophoresis showing PCR-RFLP of BER genes (a) *OGG1 Ser³²⁶Cys* (b) *MUTYH Gln³²⁴His* (c) *XRCC1 399* (d) *XRCC1 194*

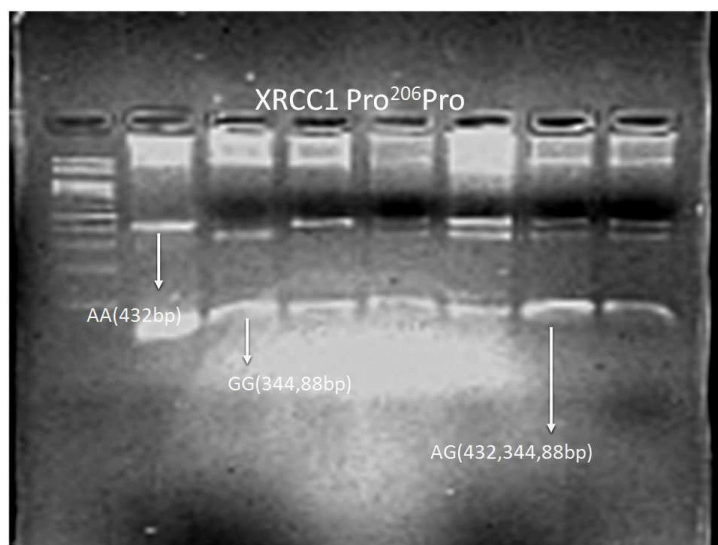
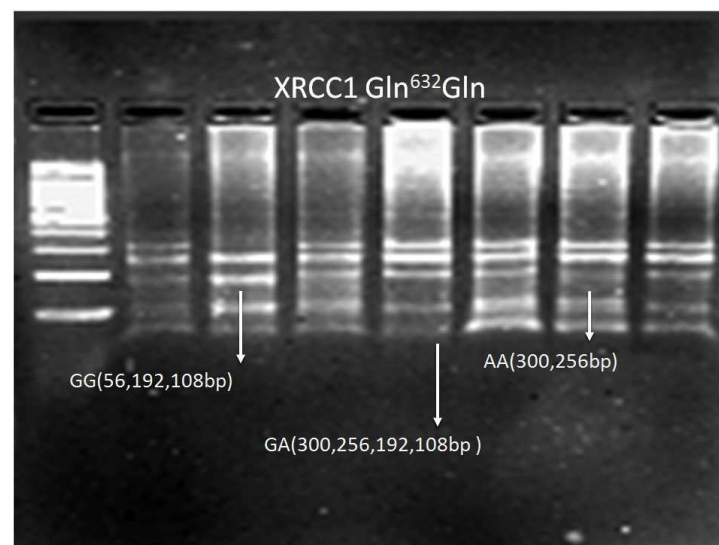
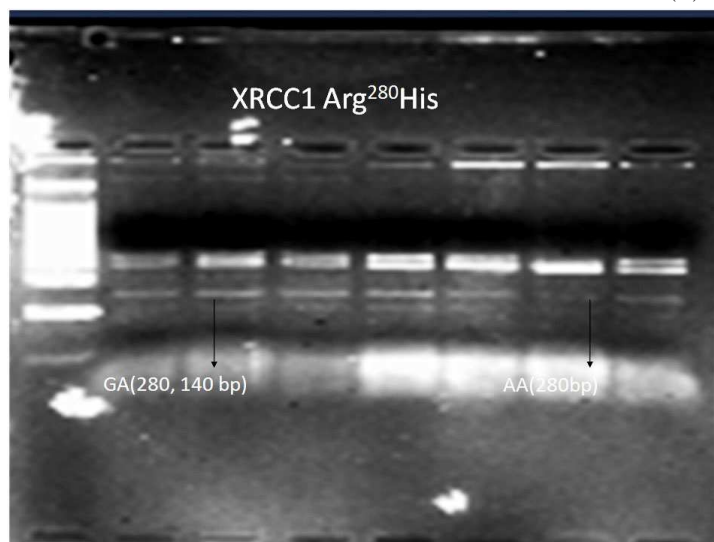
(a) *XRCC1 206*(b) *XRCC1 632*(c) *XRCC1 280*

Figure 5.3: Representative examples of agarose gel electrophoresis showing PCR-RFLP of BER genes (a) *XRCC1 206* (b) *XRCC1 632* (c) *XRCC1 194*

5.3.2 Genotyping for DSB β genes

The first polymorphic variant of the pathway studied is *XRCC1 Thr²⁴¹Met*. The polymorphism is located in exon 7 of the gene. The amplified product is of 208 bp which was digested by *NlaIII* restriction enzyme. The wild type genotype is of 208 bp, where mutant subjects was cut into two products of 120 bp and 88 bp as shown in figure 5.4a. We have also analyzed *XRCC4* gene *i.e.* Insertion/Deletion on intron 3 of *XRCC4* gene. Insertion of allele gave a product of 139 bp whereas deletion gave a product of 139 bp as indicated in figure 5.4b. For, *XRCC6 61C>G* which is located in promoter region is digested by *BanI*. The amplicon size is of 320 bp. The digestion pattern for the gene is as mentioned (CC: 262, 58 bp), (GG:180, 80, 58 bp) as represented in figure 5.4c. We also analyzed *XRCC7 G⁶⁷²¹T*. The amplified product is of 368 bp where *PvuII* is used for digestion as shown in figure 5.4d.

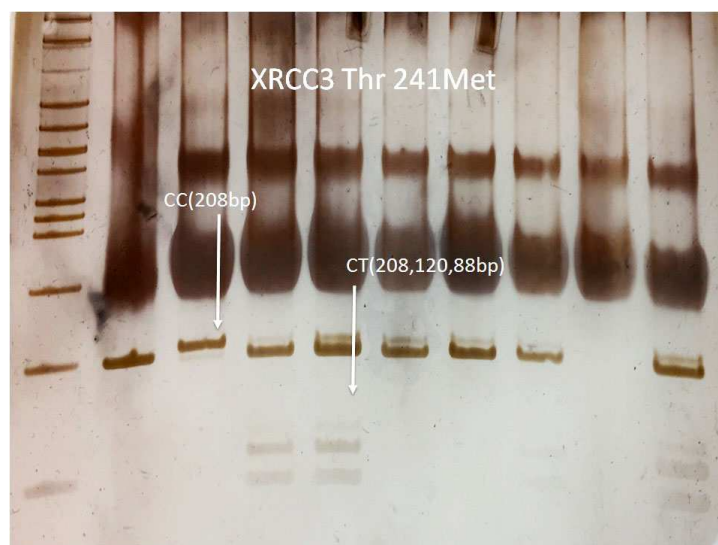
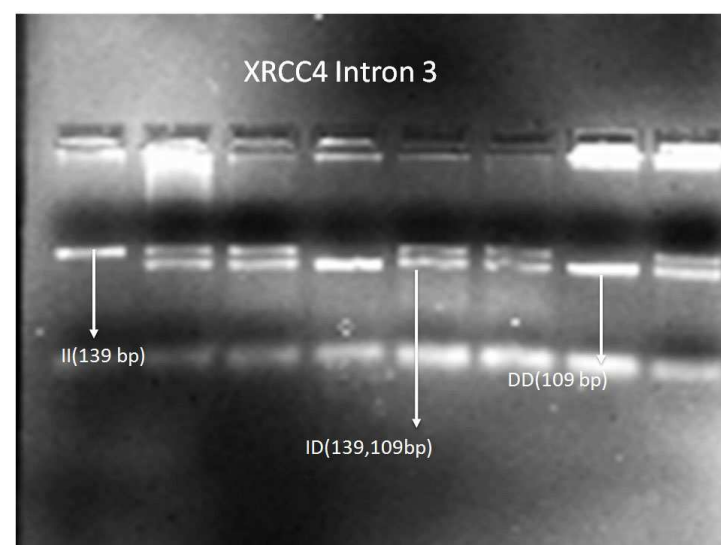
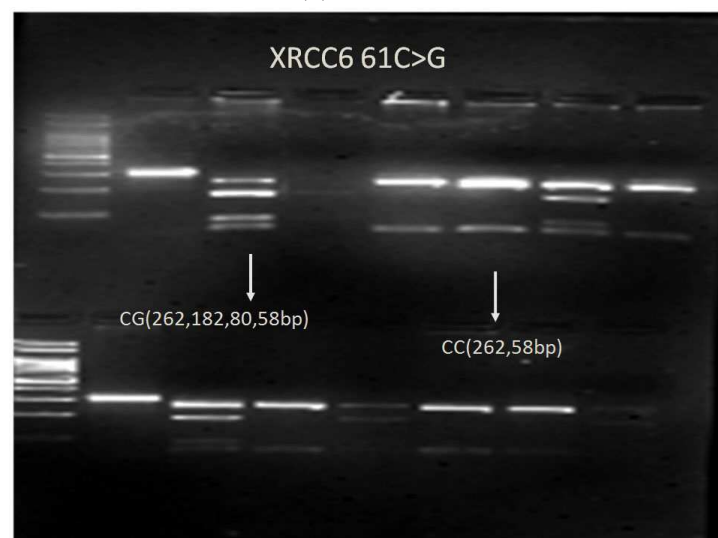
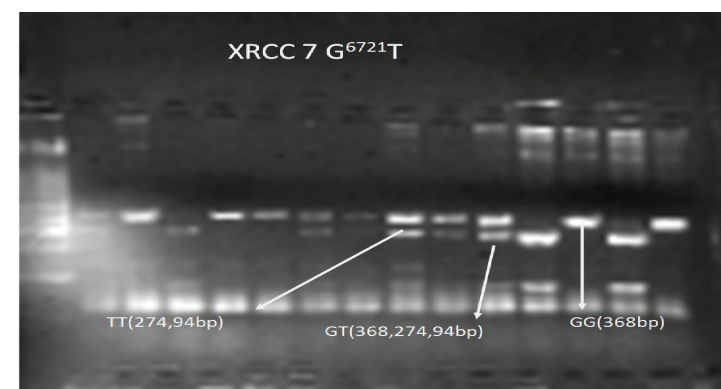
(a) *XRCC 3*(b) *XRCC 4*(c) *XRCC 6*(d) *XRCC 7*

Figure 5.4: Representative examples of agarose gel electrophoresis showing PCR-RFLP of DSB repair genes (a) *XRCC 3* (b) *XRCC 4* (c) *XRCC 6* (d) *XRCC 7*

5.4 Characteristics of the study

The present thesis work was characterized for hospital based total 358 lung cancer cases and 357 healthy controls. The demographic details of the study is presented below.

5.4.1 Age based distribution of the study group

Age is one of the most important demographic characteristic of any case-control study as it adds to the causative factor of lung cancer. The details regarding age were collected from the detailed questionnaire that was filled by patients or their relatives visiting lung cancer clinic and department of Pulmonary medicine, PGIMER, Chandigarh. The mean age of cases was 57.80 ± 10.6 years with a range of 28-86 years. We evaluated the median age for cases as 59.0 years. The mean age of controls was 53.43 ± 10.0 years with range 30-83 years, 53.0 was median age for control group. Age along with sex acted as a potent risk, therefore we evaluated the age distribution according to both sexes. 59 years being the median age for the diseased group. There was higher representation of males in patients group with age >59 years compared to the other group (90.2 vs 82.0%). Whereas there were higher number of females in age group ≤ 59 (17.9 vs 9.7%). Figure 5.5 represents the distribution.

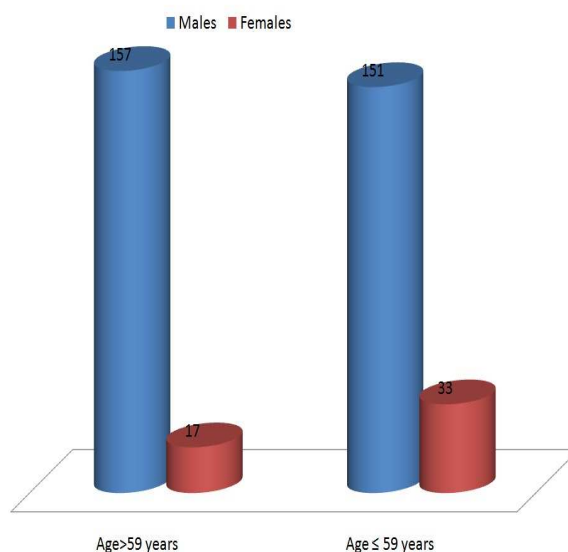


Figure 5.5: Age wise distribution of the two genders

5.4.2 Distribution of gender of the study group

The demographic details of the study was elaborated further for the distribution of gender in the study. Chi square test was performed for computing the significance of distribution of the categorical data. The study was matched for equal distribution of males and females in cases and healthy controls of the study. The number of males and females in the case group of the study were 308 (86.0%) and 50 (13.9%) respectively. Controls of the study group had 298 (83.4%) males and 59 (16.5%) females respectively. P-value for the distribution suggested no significant difference in the distribution of males and females between patients and controls, showing adequate matching ($p=0.39$) (figure 5.6).

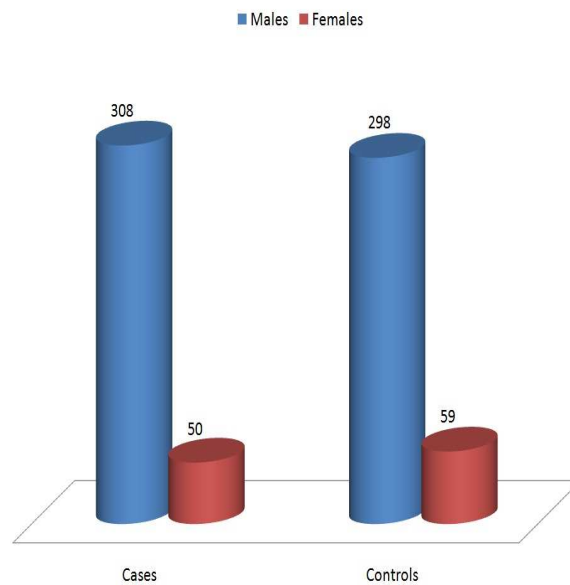


Figure 5.6: Gender based distribution of the study

5.4.3 Distribution of smoking status of the study group

Smoking is considered as the main causative factor towards lung cancer development and hence was another categorical data being evaluated for cases [figure 5.7] and controls [figure 5.8]. There were 284 (79.3%) smokers and 74 (20.6%) non-smokers in cases indicating a higher representation of smokers, suggesting smoking as a main causative factor towards lung cancer development. Lung cancer cases were matched for smoking status with controls to avoid the sample biasedness. In this respect there were 253 (70.8%) smokers and 104 (29.1%) non-smokers in the controls of the study group. Smoking status was quantified further by evaluating pack years for smokers. The mean pack years for

cases was 27.9 ± 34.6 and 17.4 ± 19.8 being for controls indicating a significant relationship between smoking and lung cancer ($p < 0.0001$). Again the pack years were analysed further for the two gender separately. It was observed that smokers who were male in both cases (87.9%) and controls (99.6%) outnumbered females both in cases (12.0%) and controls (0.4%) respectively.

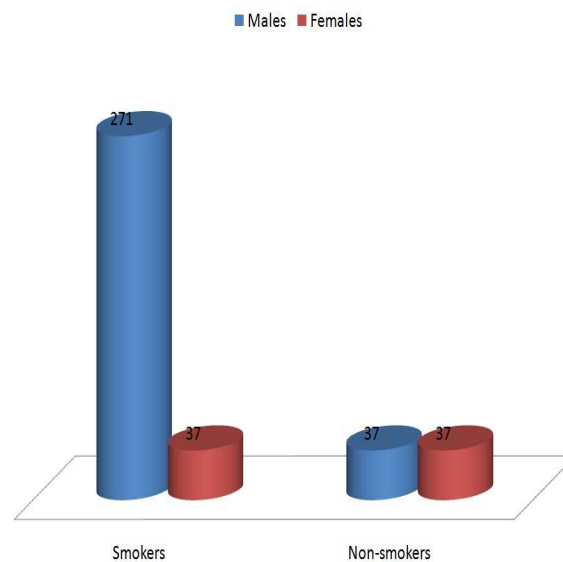


Figure 5.7: Smoking wise distribution of the two genders in cases

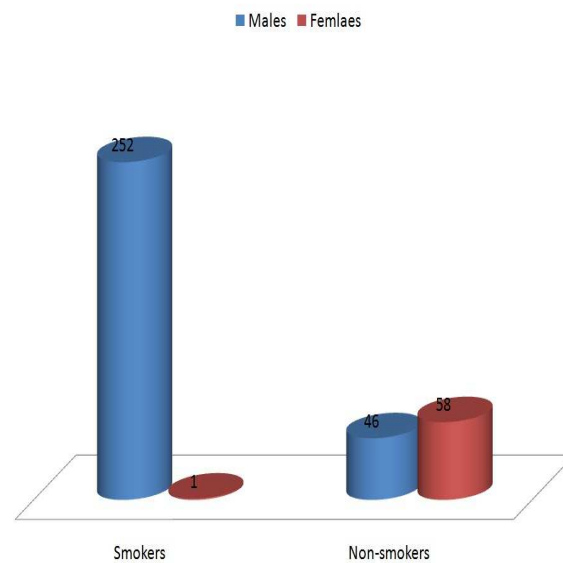


Figure 5.8: Smoking wise distribution of the two genders in controls

5.4.4 Distribution of Lung cancer patients on the basis of clinicopathological features of the study group:

The demographic details for clinicopathological features of the study were evaluated in the present thesis work. The study was classified into stage I, II, III and IV respectively. As lung cancer is majorally diagnosed at later stages of the disease in India there was higher proportion of stage III and stage IV cases compared to stage I and stage II as shown in the figure 5.9. There were 1.1% (4) stage I, 3.9% (14) stage II, 46.3% (166) stage III and 42.7% (153) stage IV lung cancer subjects. The lung cancer subjects were further stratified for the tumor extension (T), nodes involved (N) and metastasis (M). Among the lung cancer cases 8.3% (30) were Tx+T1, 12% (45) were T2, 24.8% (89) were T3 and 46.3% (166) were T4 subjects. The lymph node involvement showed the distribution as: 15.6% (56) NX+N0, 10.6% (38) N1, 38.8% (139) N2, 25.9% (93) N3 and only 1.1% with (4) N4 statistics. On the same lines metastatic status of patients is distributed as 51.9% (186) having no distant metastasis M0 and 40.2% (144) having metastasis (M1), (7.8%) or 28 of lung cancer patients had unknown TNM.

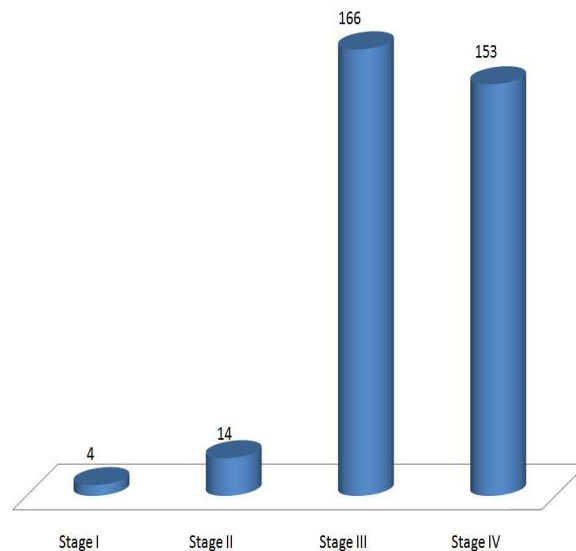


Figure 5.9: Stage based distribution of the study subjects

5.4.5 Distribution of Lung cancer study subjects on the basis of histology

Lung cancer is characterised by two major histologies which are NSCLC which includes Adenocarcinoma (ADCC) and Squamous cell carcinoma (SQCC). Whereas on the other histology is Small cell carcinoma (SCLC). Lung cancer patients in the present study were also classified accordingly. The frequency distribution classified there were 32.9% (118) patients diagnosed with ADCC subtype, 41.8% (150) having SQCC subtype, and 24.0% (86) diagnosed with SCLC subtype. The distribution analysis indicated that among the three histologies SQCC was found to be more prevalent and which had twice more patients than that of diagnosed with SCLC subtype as shown in figure 5.10. Remaining 0.8% (3) were observed to be of other histologies such as large cell carcinoma and mesothelia.

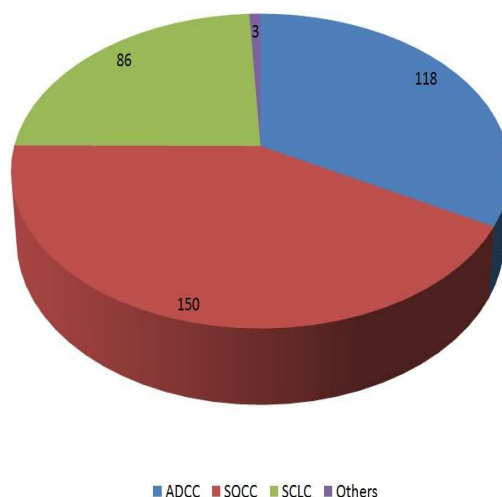


Figure 5.10: Histological distribution of Lung cancer cases

5.4.6 Distribution of Lung cancer study subjects on the basis of ECOG and KPS

KPS-PS (Karnofsky performance status) and ECOG-PS (Eastern Cooperative Oncology Group) are two important performance parameters for lung cancer patients undergoing treatment. The prognosis for treatment is made after evaluating the two performance

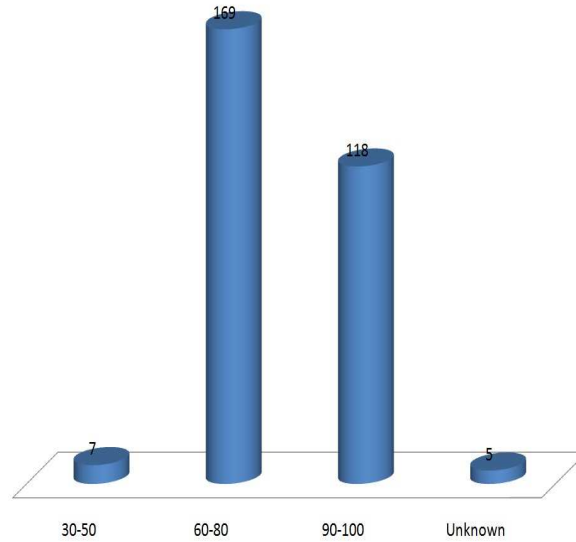


Figure 5.11: KPS distribution of Lung cancer cases

parameters. The lung cancer cases for the present study are classified according to different KPS and ECOG classes. The distribution according to the KPS suggested there were 2.3% (7) patients that showed a KPS score of 30-50, though it was least of all the three groups. Next 56.5% (169) of study subjects had KPS score between 60-80, and 39.4% (118) showed the KPS score between 90-100. While remaining 1.6% (5) did not show any KPS score (figure 5.11). With regard to ECOG-PS the study was classified as 48.4% (145) subjects with ECOG value of 0-1, 35.1% (105) with ECOG of 2 and 14.7% (44) with ECOG score between 3-4 (figure 5.12).

5.4.7 Distribution of study subjects based on Chemotherapeutic regimens administered

The most preferred treatment for lung cancer is platinum based doublet chemotherapy. Platinum drugs such as cisplatin/carboplatin along with other chemotherapeutic drugs are given to the lung cancer patients in a cyclic way through intravenous manner. The drugs given to the patients in combination with cisplatin/carboplatin are either of docetaxel, paclitaxel, irinotecan, gemcitabine and pemetrexed. The distribution of lung cancer patients according to the drug regimens administered to the patients is as follows. Out of the total subjects who underwent chemotherapeutic treatment 33.6% (76) of lung cancer subjects were given docetaxel with cisplatin/carboplatin followed by 30.5% (59) of subjects were administered irinotecan and 26.1% (69) were given pemetrexed. Remaining

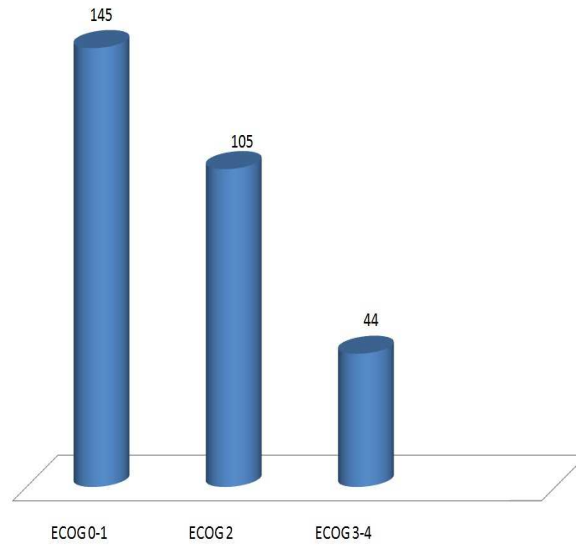


Figure 5.12: ECOG distribution of Lung cancer cases

very small number of subjects *i.e.* 5.7% (13) were administered paclitaxel followed by 0.8% (2) who were given gemcitabine along with cisplatin/carboplatin. There were 3.0% (7) lung cancer subjects with no data on drugs administered to them (figure 5.13).

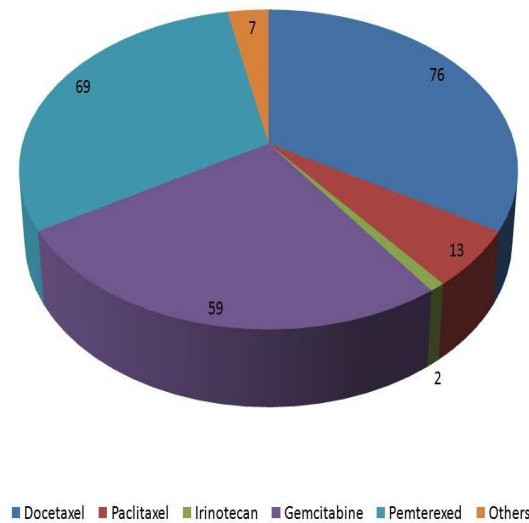


Figure 5.13: Distribution of Lung cancer subjects according to drug regimens administered

5.5 Role of polymorphic sites of BER genes *OGG1*, *MUTYH* in lung cancer susceptibility

5.5.1 Distribution of study subjects with polymorphic *OGG1* and *MUTYH* gene

We analyzed 330 healthy controls and 326 lung cancer cases for the two SNPs that are, *OGG1 Ser³²⁶Cys* and *MUTYH Gln³²⁴His* as detailed in table 5.90. The mean age for cases and control was 57.72 ± 10.57 and 53.60 ± 10.17 respectively. There was no difference in gender distribution ($p=0.305$) between cases (males: (282), 86.5%; females: (44), 13.4%) and controls (males: (275), 83.3%; females: (55), 16.6%). When smoking was taken as a factor it was observed that lung cancer patients had more representation of smokers as compared to control subjects (79.4 vs 72.1%), however the frequency of non-smokers was higher in control subjects as compared to cases (27.8 vs 20.5%). When total cumulative smoking dose was evaluated, it was observed that cases had higher mean pack years 27.9 ± 34.8 as compared to controls 17.9 ± 19.8 ($p < 0.0001$). Furthermore, when lung cancer patients were stratified on the basis of histology, it was revealed that there were 34.0% (111) adenocarcinomas, 42.3% (138) squamous cell carcinomas, 23.0% (75) small cell lung carcinomas patients.

The study was grouped according to lung cancer staging and was found that 1.2% (4) subjects were diagnosed with stage I tumor, 4.2% (14) stage II, 46.3% (151) were diagnosed with stage III patients, 42.3% (138) were stage IV patients and 5.8% (19) were unclassified. Besides stage other clinical factors included in the study are tumor size (T), lymph node involvement(N) and metastasis(M). The TNM data is available for 307 subjects. Classification according to primary tumor extension suggested 3.0% (10) were Tx, 4.6% (15) were T1, 12.2% (40) were T2, 25.7% (84) were T3, 48.4% (158) were T4 and 5.8% (19) subjects were with unknown tumor size. Data on lymph node involvement suggested there were 16.5% (54) subjects without involvement of lymph nodes in tumor progression (*i.e.*) N0, and remaining N1 to N3 had lymph nodes involved in lung carcinogenesis. In line with this there were 11.6% (38) N1 subjects, 36.5% (119) N2, 29.4% (96) N3 patients and 5.8% (19) with unknown lymph node. Data on metastasis revealed there were 41.1% (135) subjects with distant metastasis of tumor M1 compared to 52.7% (172) subjects who did not show metastasis (*i.e.*) M0.

Table 5.1: Demographic Characteristics among cases and controls

Variable	Cases N=326 (n%)	Controls N=330 (n%)	p ^a
Age(Years) Mean±SD Range	57.72±10.57	53.60±10.17	<0.0001
Gender			
Male	282 (86.5)	275 (83.3)	0.305
Female	44 (13.4)	55 (16.6)	
Smoking status			
Smokers	259 (79.4)	238 (72.1)	0.03
Non-smokers	67 (20.5)	92 (27.8)	
Pack Years Mean±SD	27.9±34.8	17.9±19.8	<0.0001
Histology			
ADCC	111 (34.0)		
SQCC	138 (42.3)		
SCLC	75 (23.0)		
Others	2 (0.6)		
TNM staging			
I	4 (1.2)		
II	14 (4.2)		
III	151 (46.3)		
IV	138 (42.3)		
Unclassified	19 (5.8)		
Tumor			
Tx	10 (3.06)		
T1	15 (4.6)		
T2	40 (12.2)		
T3	84 (25.7)		
T4	158 (48.4)		
Unknown	19 (5.8)		
Lymph node			
N0	54 (16.5)		
N1	38 (11.6)		
N2	119 (36.5)		
N3	96 (29.4)		
Unknown	19 (5.8)		
Metastasis			
M0	172 (52.7)		
M1	135 (41.4)		
Unknown	19 (5.8)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects
(a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age.
All p-values are two sided p<0.05 was considered statistically significant

5.5.2 Polymorphic *OGG1*, *MUTYH* and their role in modifying lung cancer susceptibility-Overall

The allelic, genotypic distribution and risk associated with *OGG1 Ser³²⁶Cys* and *MUTYH Gln³²⁴His* polymorphisms are summarized in table 5.2. Controls for *OGG1 Ser³²⁶Cys* showed deviation from HWE ($\chi^2=7.43$, $p=0.006$) however the lung cancer samples obeyed the HWE ($\chi^2=140.7$, $p=0.000001$). When the genotypic frequencies of *OGG1 Ser³²⁶Cys* were evaluated, it was observed that among controls there were 18.1% (60) subjects with wild (*CC*) type genotype, 57.2% (189) heterozygous (*CG*) subjects and 24.5% (81) subjects with mutant (*GG*) genotype. On the other hand when genotypic frequencies were evaluated in lung cancer cases, it was observed that there were 9.5% (31) patients with wild type (*CC*) genotype, 82.8% (270) heterozygous (*CG*) and 7.6% (25) patients had mutant (*GG*) genotype in the diseased subjects of the study group. Therefore, the analysis revealed that in case of *OGG1 Ser³²⁶Cys* both the wild (*CC*) and mutant (*GG*) genotype were found to be over-represented in controls as compared to cases. A significant difference between genotypic frequencies of cases and controls was observed ($\chi^2=53.0$, $df=2$, $p<0.0001$). The minor allele frequency for the controls was 0.48 as compared to 0.49 for cases. As cases reported a high frequency of single allelic carriers (*CG*) compared to controls (82.8 vs 57.2%), logistic regression analysis revealed that individuals who were single allelic variant (*CG*) carriers had a two-fold increased risk towards lung cancer (OR=2.44, 95%C.I.=1.50-3.9, $p=0.0003$) when compared to the group carrying the wild (*CC*) type genotype. On the contrary carriers of double allelic variants (*GG*) or mutants exhibited a protective effect towards lung cancer (OR=0.4, 95%C.I.=0.22-0.87, $p=0.0185$). Combining both the single (*CG*) and double allelic variants (*GG*) as a single genotype our data revealed a two-fold risk towards lung cancer (OR=1.88, 95%C.I.=1.17-3.05, $p=0.009$) which was found to be significant as shown in table 5.2.

MUTYH Gln³²⁴His polymorphism was the next SNP evaluated in the study. The genotypic frequencies for controls of *MUTYH Gln³²⁴His* followed the Hardy Weinberg Equilibrium (HWE) ($\chi^2=1.13$, $p=0.286$). It was observed that controls had a higher frequency of wild (*GG*) (68.7% vs 48.7%) and mutant genotype (*CC*) (3.9% vs 1.5%) as compared to cases respectively. There was a significant difference between the genotypic frequencies of cases and controls ($\chi^2=36.0$, $df=2$, $p<0.0001$). Minor allele frequency was higher in cases (0.26) as compared to controls (0.17). Subjects carrying single allelic carriers (*GC*) or heterozygous genotype within controls (49.6%) were significantly

over-represented in comparison to cases (27.2%) as shown in table 5.2. Henceforth, heterozygous (*GC*) subjects were observed to acquire two-fold risk towards lung cancer (OR=2.6, 95%C.I.=1.8-3.7, $p<0.0001$) as compared to individuals carrying both the alleles of wild type genotype. Subjects carrying mutant genotype (*CC*) did not show any effect towards lung cancer (OR=0.59, 95%C.I.=0.20-1.75, $p=0.34$), whereas subjects with combined variant genotype (*GC+CC*) indicated a high risk towards lung cancer development (OR=2.64, 95%C.I.=1.88-3.70, $p<0.0001$).

Table 5.2: Assessment of *OGG1* and *MUTYH* risk towards Lung cancer susceptibility

OGG1 Ser³²⁶ Cys (C/G) rs 1052133	Controls N=330 (n%)	Cases N=326 (n%)	Overall AOR (95% CI)^a	p^b
CC	60 (18.1)	31 (9.5)	1	
CG	189 (57.2)	270 (82.8)	2.44 (1.50-3.9)	0.0003
GG	81 (24.5)	25 (7.6)	0.4 (0.22-0.87)	0.0185
CG+GG	270 (81.8)	295 (90.4)	1.88 (1.17-3.05)	0.0093
C	309	332		
G	351	320		
MAF	0.48	0.49		
MUTYH Gln³²⁴ His (G/C) rs 3219489	Controls N=330 (n%)	Cases N=326 (n%)	Overall AOR (95% CI)^b	p^a
GG	227 (68.7)	159 (48.7)	1	
GC	90 (27.2)	162 (49.6)	2.6 (1.8-3.70)	<0.0001
CC	13 (3.9)	5 (1.5)	0.59 (0.20-1.75)	0.34
GC+CC	103 (31.2)	167 (51.2)	2.64 (1.88-3.70)	<0.0001
G	544	480		
C	116	172		
MAF	0.17	0.26		

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.5.3 Genotypic and allelic distribution of polymorphic *OGG1*, *MUTYH* gene and lung cancer susceptibility based on histology

Table 5.3 shows the association of both *OGG1* Ser³²⁶Cys and *MUTYH* Gln³²⁴His polymorphic sites towards lung cancer risk after stratifying subjects according to histology. In case of *OGG1* Ser³²⁶Cys for ADCC histological subtype it was observed that frequency

of single allelic carriers (CG) was 86.4% as compared to 57.2% in controls and when compared to wild type genotype (CC) a 4.7 fold risk towards ADCC development (OR=4.72, 95%C.I.=2.0-10.65, $p=0.0002$) was observed. On the contrary, subjects with double allelic variant genotype (GG) or mutant genotype showed no association towards lung cancer risk (OR=0.67, 95%C.I.=0.19-2.27, $p=0.52$). However, when both single allelic variants and mutant genotypes were combined as a single genotype ($CG+GG$) and risk was evaluated after adjusting for various factors like age, sex and smoking, it was observed that subjects possessed a almost four-fold risk (OR=3.63, 95%C.I.=1.61-8.16, $p=0.0018$) towards lung cancer.

In case of SQCC a higher frequency of subjects with mutant genotype (GG) was observed in controls (6.5%) as compared to cases (24.5%) and after applying logistic regression the odds ratio suggested a protective effect towards lung cancer development (OR=0.26, 95%C.I.=0.10-0.65, $p=0.04$) when compared to wild type genotype (CC) (table 5.3). Furthermore, subjects diagnosed with SCLC subtype and carrying single allelic variant genotype or heterozygous genotype (CG) exhibited an increased frequency for cases (81.3%) as compared to controls (57.2%). Taking wild type genotype (CC) as a reference a high risk towards SCLC (OR=2.55, 95%C.I.=0.98-6.61, $p=0.052$) was observed. The frequency of mutant genotype (GG) was observed to be higher in controls (24.5%) as compared to SCLC subjects (10.6%) thereby a decreased risk of developing lung cancer was observed in such subjects though it was not significant as shown in table 5.3.

The stratified analysis for the polymorphic variant *MUTYH Gln³²⁴His* exhibited that the frequency distribution analysis classified 54.9% subjects with wild type (GG) genotype, 43.2% with heterozygous (GC) and 1.8% subjects with mutant (CC) genotype as shown in table 5.3. After adjusting for factors such as smoking, age and gender using logistic regression analysis it was revealed that subjects with ADCC subtype and carrying single allelic variant genotype (GC) showed a two fold risk towards lung cancer (OR=2.08, 95%C.I.=1.30-3.3, $p=0.002$) as compared to subjects with wild type genotype (GG). In contrast to it, ADCC subjects having double allelic variant genotype (CC) showed a trend of protective effect for lung cancer development (OR=0.59, 95%C.I.=0.12-2.85, $p=0.51$). Furthermore, when we combined both the heterozygous and mutant genotype (*i.e.*) ($GC+CC$) as a single genotype it was observed that ADCC subgroup showed a higher frequency (45.0%) compared to (31.2%) for controls. Furthermore, it was also observed that subjects diagnosed with ADCC subtype acquired a high risk towards lung cancer for combined variant genotype (OR=1.88, 95%C.I.= 1.19-2.98, $p=0.006$) on comparison with wild type subjects.

The frequency distribution for SQCC and SCLC subtypes stated that there were 54.3

and 50.6% single allelic carriers (GC) respectively as compared to 27.2% for controls. An increase in risk towards both subtypes of lung cancer for the heterozygous genotype (GC) was observed in *i.e.* SQCC (OR=2.88, 95%C.I.=1.86-4.44, $p<0.0001$) and SCLC patients (OR=2.3, 95%C.I.=1.34-4.14, $p=0.002$). The analysis when followed for the mutant genotype (CC) for SQCC subtype indicated a protective affect (OR=0.26, 95%C.I.= 0.03-2.24, $p=0.22$). Furthermore when both mutant (CC) and heterozygous (GC) genotype were combined as a single genotype ($GC+CC$) a significant risk was observed (OR=2.5, 95%C.I.=1.67-3.9, $p<0.0001$) on comparison with the wild (GG) type genotype. Similarly, lung cancer subjects diagnosed with SCLC subtype and carrying combined variant genotype ($GC+CC$) were highly represented as compared to controls (53.3 vs 31.2%) . On comparing with wild type genotype (GG) risk analysis depicted a 2-fold propensity for lung cancer development (OR=2.26, 95%C.I.=1.3-3.89, $p<0.003$).

Table 5.3: Polymorphic *OGG1* and *MUTYH* towards lung cancer development based on Histology

Genes OGG1 Ser³²⁶ Cys (C/G)	Controls N=330 (n%)	ADCC			SQCC			SCLC		
		Cases N=111 (n%)	AOR (95% CI) ^a	p ^b	Cases N=138 (n%)	AOR (95% CI) ^a	p ^b	Cases N=75 (n%)	AOR (95% CI) ^a	p ^b
CC	60 (18.1)	8 (7.2)	1		17 (12.3)	1		6 (8.0)	1	
CG	189 (57.2)	96 (86.4)	4.72 (2.0-10.65)	0.0002	112 (81.1)	1.32 (0.70-2.49)	0.38	61 (81.3)	2.55 (0.98-6.61)	0.052
GG	81 (24.5)	7 (6.3)	0.67 (0.19-2.27)	0.52	9 (6.5)	0.26 (0.10-0.65)	0.004	8 (10.6)	0.43 (0.11-1.62)	0.21
CG+GG	270 (81.8)	103 (92.7)	3.63 (1.61-8.16)	0.0018	121 (87.6)	1.0 (0.54-1.89)	0.97	69 (92.0)	1.98 (0.76-15.1)	0.158
MUTYH Gln³²⁴ His (G/C)	Controls N=330 (n%)	Cases N=111 (n%)	AOR (95% CI) ^a	p ^b	Cases N=138 (n%)	AOR (95% CI) ^a	p ^b	Cases N=75 (n%)	AOR (95% CI) ^a	p ^b
GG	227 (68.7)	61 (54.9)	1		62 (44.9)	1		35 (46.6)	1	
GC	90 (27.2)	48 (43.2)	2.08 (1.30-3.3)	0.0022	75 (54.3)	2.88 (1.86-4.44)	< 0.0001	38 (50.6)	2.3 (1.34-4.14)	0.0027
CC	13 (3.9)	2 (1.8)	0.59 (0.12-2.85)	0.51	1 (0.7)	0.26 (0.03-2.24)	0.22	2 (2.6)	1.18 (0.24-5.71)	0.83
GC+CC	103 (31.2)	50(45.0)	1.88 (1.19-2.98)	0.0065	76 (55.0)	2.5 (1.67-3.91)	< 0.0001	40 (53.3)	2.26 (1.3-3.89)	0.003

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.5.4 Effect of smoking status on association of polymorphic *OGG1*, *MUTYH* and lung cancer risk

Association analysis was extended further to investigate the effect of smoking on lung cancer in relation to the polymorphic *OGG1* and *MUTYH* genes. It was observed that among the smokers who were lung cancer patients the frequency of subjects with wild type genotype (*CC*) for *OGG1 Ser³²⁶Cys* was 9.6% whereas 82.2% were single allelic carriers (*CG*) and 8.1% subjects carried double allelic variant genotype (*GG*). Whereas, controls reported a higher frequency of carriers for both wild alleles (*CC*) and mutant (*GG*) alleles. Among the controls subjects who were smokers, 11.3% had wild type genotype (*CC*), 62.1% carried heterozygous genotype (*CG*) and 26.4% were double allelic variant genotype (*GG*) as shown in table 5.4. Taking the wild type (*CC*) genotype as a reference group it was observed that smokers with both mutant alleles (*GG*) for *OGG1 Ser³²⁶Cys* showcased a protective effect towards lung cancer (OR=0.3, 95%C.I.=0.13-0.6, ***p*=0.002**). On the contrary subjects who were non-smokers and carried heterozygous genotype (*CG*) reported a higher risk towards lung cancer (OR=7.9, 95%C.I.=2.94-21.2, ***p*<0.0001**). Since the frequency of the mutant genotype (*GG*) was less we combined the two genotypes together as a single genotype (*CG+GG*). Association analysis for the above genotype suggested a high risk of lung cancer for non-smokers (OR=5.7, 95%C.I.=2.2-15.1, ***p*=0.0004**).

In case of *MUTYH Gln³²⁴His* within smokers group controls had a higher frequency of wild type genotype (*GG*) as compared to cases (65.9% vs 48.2%) respectively whereas heterozygous genotype (*GC*) were over-represented in cases (50.5%) as compared to the controls (30.2%). Regression analysis for the same suggested that smokers who were heterozygous (*GC*) for *MUTYH Gln³²⁴His* exhibited a two-fold risk for lung cancer (OR=2.35, 95%C.I.=1.59-3.4, ***p*=<0.0001**) as compared to wild type genotype (*GG*). Similar to the smokers, non-smokers had a higher representation of wild type (*GG*) genotype in controls as compared to cases (76.0% vs 50.7%) whereas, single allelic (*GC*) carriers were higher in cases compared to controls (46.2% vs 19.5%). Therefore, subjects harbouring heterozygous genotype (*GC*) within non-smokers group were at three-fold risk (OR= 3.37, 95%C.I.=1.62-7.02, ***p*=0.001**) towards lung cancer as shown in table 5.4.

Table 5.4: Genotypic distribution of OGG1 and MUTYH based on smoking status and its association with risk of Lung cancer

Genes	Smokers				Non Smokers			
OGG1 Ser³²⁶ Cys (C/G)	Controls N=238 (n%)	Cases N=259 (n%)	AOR (95% CI)^a	p^b	Controls N=92 (n%)	Cases N=67 (n%)	AOR (95% CI)^a	p^b
CC	27 (11.3)	25 (9.6)	1.0		33 (35.8)	6 (8.9)	1	
CG	148 (62.1)	213 (82.2)	1.41 (0.79-2.58)	0.25	41 (44.5)	57 (85.0)	7.9 (2.94-21.2)	< 0.0001
GG	63 (26.4)	21 (8.1)	0.3 (0.13-0.6)	0.002	18 (19.5)	4 (5.9)	0.77 (0.16-3.78)	0.75
CG+GG	211 (88.6)	234 (90.3)	1.04 (0.57-1.90)	0.87	59 (64.1)	61 (91.0)	5.7 (2.2-15.1)	0.0004
MUTYH Gln³²⁴ His (G/C)	Controls N=238 (n%)	Cases N=259 (n%)	AOR (95% CI)^a	p^b	Controls N=92 (n%)	Cases N=67 (n%)	AOR (95% CI)^a	p^b
GG	157 (65.9)	125 (48.2)	1.0		70 (76.0)	34 (50.7)	1	
GC	72 (30.2)	131 (50.5)	2.35 (1.59-3.4)	< 0.0001	18 (19.5)	31 (46.2)	3.37 (1.62-7.02)	0.001
CC	9 (3.7)	3 (1.15)	0.48 (0.2-1.93)	0.30	4 (4.3)	2 (2.9)	0.95 (0.16-5.58)	0.95
GC+CC	81 (34.0)	134 (51.7)	2.14 (1.47-3.12)	0.0001	22 (23.9)	33 (49.2)	2.91 (1.45-5.80)	0.0025

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender.(b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.5.5 Relation between clinicopathological characteristics and polymorphic *OGG1* and *MUTYH* distribution

The present study was evaluated further for clinical stage, primary tumor extension and metastatic state and its association with the polymorphic variants of both *OGG1* and *MUTYH*. Table 5.5 shows the distribution of genotype according to various clinicopathological features and their relation towards lung cancer. The clinical stage distribution divided the study into stage III and stage IV patients. On the basis of primary tumor extension the study was stratified into TX+T1+T2 and T3+T4 subsets. Metastasis divided the study into two subgroups M0 and M1.

The genotypic frequencies were calculated for each polymorphic site for both stage III and stage IV subjects as shown in table 5.5. Genotypic frequencies were compared to find out if any significant difference was observed in the genotypic frequencies among the two subgroups of stage of lung cancer. The association analysis between stage and genotype for *OGG1 Ser³²⁶Cys* suggested no significant relation towards lung cancer as there was no significant difference between genotypes. As seen in case of *OGG1 Ser³²⁶Cys* 83.7 % of stage III subjects were single allelic carriers (*CG*) and 83.5 % of stage IV patients were of (*CG*) genotype. Logistic analysis pointed out, a visible trend towards increased risk of lung cancer for the mutant genotype (*GG*) (OR=2.24, 95%C.I.=0.66-7.51, p=0.19). In case of *MUTYH Gln³²⁴His* the frequency distribution for heterozygous subjects (*GC*) was observed to be similar in both stage III and stage IV subjects (49.4% vs 48.5%) respectively. Logistic regression analysis exhibited no association with lung cancer as indicated in table 5.5 when compared to wild type subjects (OR=0.96, 95%C.I.=0.60-1.53, p=0.87). Similarly combined variants (*GC+CC*) were similar in frequency (50.5 vs 50.5%) for both stage IV and stage III subjects.

Tumor progression was elaborated further in detail by analysing the correlation between genotype and primary tumor extension. It was observed that lung cancer subjects diagnosed with extended tumor size T3+T4 (242) were represented more as compared to patients with TX+T1+T2 tumors (65). There was no significant difference in the frequency distribution of combined variants (*CG+GG*) with (T3+T4) tumor size as compared to subjects with TX+T1+T2 tumor indicating no correlation with tumor extension.

In case of *MUTYH Gln³²⁴His* frequency of combined variant (*GC+CC*) genotype in subjects with T3+T4 was higher (48.7%) as compared to subjects with TX+T1+T2 category (46.1%). However, primary tumor extension was not observed to modulate the lung cancer risk as indicated in table 5.5.

Table 5.6 represents the distribution of genotypic frequencies of subjects and their relationship with lymph node involvement. Subjects carrying single allelic variant genotype (*CG*) and under N1+N2+N3 category showed a higher prevalence (89.3%) as compared to subjects under N0 category (74.0%). Odds ratio exhibited a three fold risk towards lung cancer (OR=3.0, 95%C.I.=1.29-7.16, ***p*=0.01**) indicating subjects carrying heterozygous genotype had a higher probability for lymph node invasion. Following the trend, subjects carrying the combined variant (*CG+GG*) genotype revealed lymph node involvement in subjects with lung cancer outcome (OR=2.9, 95%C.I.=1.26-6.9, ***p*=0.012**).

Relationship of different genotypes of both *OGG1* and *MUTYH* with metastatic status is demonstrated in table 5.6. In case of *OGG1 Ser³²⁶Cys* the genotypic frequencies were evaluated for the two groups and comparison showed there were 9.6% subjects that showed metastasis and 9.3% that did not show metastasis (M0) for lung cancer patients carrying the wild type genotype (*CC*). The frequency of subjects harboring single allelic variant genotype (*CG*) was 82.2% for subjects with distant metastatic tumor as compared to subjects lacking metastasis (84.8%). The combined variants (*CG+GG*) were also similar in frequency for the two metastatic and non-metastatic groups (90.3% vs 90.6%). Since there was no significant difference in the frequency distribution among the two above discussed genotypic groups, it was evident that this polymorphic site did not show any correlation with metastasis or spread of lung cancer to other organs. In case of *MUTYH Gln³²⁴His* the proportion of heterozygous subjects (*GC*) with distant metastasis was 49.6% compared to 47.0% subjects without metastasis. However, there was no significant difference observed in the frequency distribution of subjects. Similar to the previous polymorphic variant *MUTYH* was observed not to contribute towards the progression of lung cancer among other organs.

Table 5.5: Relation between clinicopathological factors and *OGG1* and *MUTYH* (Stage and primary tumor extension)

Genotype	Clinical stage		AOR (95% C.I.) ^a	p ^b	Primary tumor extension		AOR (95% C.I.) ^b	p ^b
	III	IV			Tx+ T1+ T2	T3+ T4		
OGG1 Ser ³²⁶ Cys (C/G)	172	140			65	242		
CC	18 (10.4)	11 (7.8)	1		5 (7.6)	24 (9.9)	1	
CG	144 (83.7)	117 (83.5)	1.29 (0.57-2.92)	0.53	56 (86.1)	201 (83.0)	0.78 (0.28-2.6)	0.65
GG	10 (5.8)	12 (8.5)	2.24 (0.66-7.51)	0.19	4 (6.1)	17 (7.0)	1.5 (0.26-8.5)	0.64
CG+GG	154 (89.5)	129 (92.1)	1.35 (0.60-3.05)	0.46	60 (92.3)	218 (90.0)	0.81 (0.29-2.2)	0.68
MUTYH Gln ³²⁴ His (G/C)	172	140			65	242		
GG	85 (49.4)	70 (50.0)	1		35 (53.8)	120 (49.5)	1	
GC	85 (49.4)	68 (48.5)	0.96 (0.60-1.53)	0.87	30 (46.1)	118 (48.7)	1.12 (0.64-1.96)	0.66
CC	2 (1.16)	2 (1.42)	0.67 (0.08-5.67)	0.72	-	4 (1.7)		
GC+CC	87 (50.5)	70 (50.0)	0.95 (0.60-1.51)	0.84	30 (46.1)	122 (50.4)	1.17 (0.67-2.0)	0.57

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding (b)-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

Table 5.6: Relation between clinicopathological factors and *OGG1* and *MUTYH* (Lymph node and metastasis)

Genotype	Lymph node invasion		AOR (95% C.I.) ^a	p ^b	Metastasis		AOR (95% C.I.) ^a	p ^b
	N0	N1+ N2+ N3			M0	M1		
OGG1 Ser ³²⁶ Cys (C/G)	54	253			172	135		
CC	10 (18.5)	19 (7.8)	1		16 (9.3)	13 (9.6)	1	
CG	40 (74.0)	217 (89.3)	3.0 (1.29-7.16)	0.01	146 (84.8)	111 (82.2)	0.8 (0.4-1.9)	0.79
GG	4 (7.4)	17 (6.9)	2.0 (0.48-8.4)	0.33	10 (5.8)	11 (8.1)	1.49 (0.45-4.9)	0.5
CG+GG	44 (81.4)	234 (96.2)	2.9 (1.26-6.9)	0.012	156 (90.6)	122 (90.3)	0.93 (0.42-2.0)	0.85
MUTYH Gln ³²⁴ His (G/C)	54	253			172	135		
GG	29 (53.7)	126 (51.8)	1		89 (51.7)	66 (48.8)	1	
GC	24 (44.4)	124 (51.0)	1.17 (0.64-2.13)	0.68	81 (47.0)	67 (49.6)	1.02 (0.70-1.80)	0.62
CC	1 (1.8)	3 (1.2)	1.0 (0.09-11.6)	0.96	2 (1.1)	2 (1.4)	0.68 (0.11-4.20)	0.68
GC+CC	25(46.2)	127 (52.2)	1.15 (0.64-2.10)	0.62	83 (48.2)	69 (51.1)	1.05 (0.66-1.67)	0.82

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding (b)-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

5.5.6 Combinatorial assessment of polymorphic *OGG1* and *MUTYH* towards lung cancer predisposition

The combinatorial assessment of the SNPs of both *OGG1 Ser³²⁶Cys* and *MUTYH Gln³²⁴His* and their synergistic role towards development for lung cancer is assessed in table 5.7. Our data showed that the frequency of wild type (*CC+GG*) genotype was found to be more represented in control (43.6%) subjects as compared to cases (9.74%) for the combination of both *OGG1* and *MUTYH* SNP's. Single allelic carriers (*CG+GC*) were higher in cases (89.5%) as compared to controls (53.1%). Regression analysis revealed that subjects who were carrying a single allelic variant copy of both the SNPs' *i.e.* (*CG+GC*) had a very high propensity towards lung cancer *i.e.* roughly seven-fold (OR=7.26, 95%C.I.=3.5-14.97, $p<0.0001$). However, the carriers of double allelic variants (*GG+CC*) did not show any association towards lung cancer (OR=0.8, 95%C.I.=0.06-10.7, $p=0.8$). Furthermore, when the combined genotypes of both SNPs were stratified towards histology, it was observed that individuals with ADCC subtype and who carried heterozygous genotype for both the combined SNPs (*CG/+ GC*) had a higher (90.4%) frequency compared to controls (53.1%). A 14-fold risk was observed towards lung cancer as compared to subjects carrying the wild alleles for both SNPs (OR=14.1, 95%C.I.=4.00-50.30, $p=<0.001$). Subjects carrying combined variant genotype (*CG/+ GC*) had very high frequency in case of both SQCC (89.7%) and SCLC (91.1%) subtype. Association analysis exhibited a five-fold risk in case of SQCC (OR=5.0, 95%C.I.=1.92-13.10, $p=0.001$) towards lung cancer. Similarly subjects diagnosed with SCLC subtype and carrying combined variant genotype revealed an almost six-fold risk (OR=5.7, 95%C.I.=1.39-23.4, $p=0.015$) towards lung cancer as shown in table 5.8.

Table 5.7: Combinatorial assessment of OGG1 and MUTYH towards Lung cancer for overall

OGG1 Ser ³²⁶ Cys and MUTYH Gln ³²⁴ His	Controls N=94 (n%)	Overall		
		Cases N=144 (n%)	Overall AOR (95% CI) ^a	p ^b
CC/GG	41 (43.6)	14 (9.7)	1	
CG/GC	50 (53.1)	129 (89.5)	7.26 (3.5-14.9)	< 0.0001
GG/CC	3 (3.1)	1 (0.6)	0.8 (0.06-10.78)	0.8
CG+GG/GC+CC	53 (56.3)	130 (90.2)	6.8 (3.24-14.7)	< 0.0001

(a) Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b) Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

Table 5.8: Combinatorial assessment of OGG1 and MUTYH towards Lung cancer and Histology for overall

OGG1 Ser ³²⁶ Cys and MUTYH Gln ³²⁴ His	Controls N=94 (n%)	ADCC			SQCC			SCLC		
		Cases N=42 (n%)	Overall AOR (95% CI) ^a	p ^b	Cases N=68 (n%)	Overall AOR (95% CI) ^a	p ^b	Cases N=34 (n%)	Overall AOR (95% CI) ^a	p ^b
CC/GG	41 (43.6)	4 (9.5)	1		7 (10.2)	1		3(8.8)	1	
CG/GC	50 (53.1)	38 (90.4)	14.1 (4.0-50.3)	< 0.001	60 (88.2)	5.0 (1.92-13.0)	0.001	31 (91.1)	5.7 (1.7-23.4)	0.015
GG/CC	3 (3.1)	-	-	-	1 (1.4)	2.0 (0.13-29.30)	0.61	-	-	-
CG+GG/GC+CC	53 (56.3)	38 (90.4)	12.6 (3.60-43.50)	0.0001	61 (89.7)	4.75 (1.84-12.26)	0.0013	31 (91.1)	5.3 (1.31-21.48)	0.019

(a) Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b) Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.6 Role of polymorphic sites of *XRCC1* gene in lung cancer susceptibility

5.6.1 Demographic details of lung cancer study subjects and controls with polymorphic *XRCC1* gene

The investigative study for the five SNPs of *XRCC1* enrolled a total of 330 lung cancer cases and 325 control subjects as shown in table 5.9. The mean age for the cases and controls were 57.9 ± 10.58 and 52.27 ± 10.84 respectively. There were a total of 285 (86.3%) males and 45 (13.6%) females in the diseased group whereas, controls had a total of 265 (81.5%) males and 60 (18.4%) females respectively. The study was also patterned for the difference in the smoking status of the two study groups. A higher prevalence of smokers was represented in cases as compared to controls (81.2% vs 68.6%, $p=0.0003$). When stratified according to pack years, cases had higher mean values of pack years smoked 29.9 ± 40.9 in comparison to 17.02 ± 19.5 pack years of controls ($p < 0.0001$) as shown in table 5.9.

The lung cancer cases were further classified on different pathological parameters. Out of 330 lung cancer cases 32.1% (106) were diagnosed with adenocarcinomas, 42.1% (139) with squamous cell carcinoma and 24.5% (81) with small cell carcinoma. Among the other clinicopathological factors distribution according to stage suggested there were only 0.9% (3) stage I patients, 4.5% (15) stage II patients, 46.9% (155) stage III patients and 43.6% (144) stage IV patients whereas remaining 3.9% (13) patients were unclassified. Lung cancer study subjects were also classified on the basis of TNM staging *i.e.* primary tumor extension, lymph node involvement, and metastasis. The TNM classification revealed that there were 3.3% (11) subjects with non-assessable cancer Tx, 5.1% (17) subjects with T1 tumor extension. Furthermore, there were 12.4% (41) subjects with T2, 25.1% (83) with T3 and 48.1% (159) with T4 size of cancer tumor indicating the higher stage of cancer, on the contrary there were 5.7% (19) subjects with unclassified tumor extension. Further, when the patients were stratified according to lymph node involvement it was found that there were 15.7% (52) patients with no involvement of lymph nodes (N0+NX), whereas 11.2% (37) of lung cancer subjects were classified under N1 suggesting involvement of lymph nodes present in lungs or airways joining lungs. Furthermore, 38.7% (128) of patients were under N2 category indicating involvement of lymph nodes in the center of the chest or on the same side of the affected lung and 27.5% (91) were with N3 category highlighting involvement of lymph nodes on the opposite side of chest from affected lung

or on the top of the lung and remaining 1.2% (4) were with N4, 5.7% (19) subjects in the study had no information regarding lymph nodes involvement. Metastatic stratification showed there were 53.9% (178) of study subjects without any spread of lung cancer *i.e.* M0 and 40.3% (133) of subjects with cancer spread to other parts of body *i.e.* M1a+M1b and rest 5.7% (19) of the unclassified metastatic state as stated in table 5.9.

Table 5.9: Demographic Characteristics among cases and controls of *XRCC1*

Variable	Cases N=330 (n%)	Controls N=325 (n%)	p ^a
Age(Years) Mean±SD Range	57.9±10.58	52.27±10.84	<0.0001
Gender			
Male	285 (86.3)	265 (81.5)	0.114
Female	45 (13.6)	60 (18.4)	
Smoking status			
Smokers	268 (81.2)	223 (68.6)	0.0003
Non-smokers	62 (18.7)	102 (31.3)	
Pack Years Mean±SD	29.9±40.9	17.02±19.5	<0.0001
Histology			
ADCC	106 (32.1)		
SQCC	139 (42.1)		
SCLC	81 (24.5)		
Others	4 (1.2)		
TNM staging			
I	3 (0.9)		
II	15 (4.5)		
III	155 (46.9)		
IV	144 (43.6)		
Unclassified	13 (3.9)		
Tumor			
Tx	11 (3.3)		
T1	17 (5.1)		
T2	41 (12.4)		
T3	83 (25.1)		
T4	159 (48.1)		
Unknown	19 (5.7)		
Lymph node			
N0	52 (15.7)		
N1	37 (11.2)		
N2	128 (38.7)		
N3	91 (27.5)		
Unknown	19 (5.7)		
Metastasis			

to be cont'd on next page

M0	178 (53.9)		
M1	133 (40.3)		
Unknown	19 (5.7)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects,
(a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age.
All p-values are two sided p<0.05 was considered statistically significant

5.6.2 Genotypic and allelic frequencies of *XRCC1* variants and their association with overall lung cancer risk

The allelic and genotypic distribution of the five polymorphic variants of (*XRCC1*) is shown in table 5.10. In our study genotypic frequencies for controls were in accordance with Hardy Weinberg Equilibrium (HWE) in case of *XRCC1 Arg¹⁹⁴ Trp* ($\chi^2=0.008$, p=0.92), *Pro²⁰⁶ Pro* ($\chi^2=0.48$, p=0.48) and *Arg³⁹⁹ Gln* ($\chi^2=2.28$, p=0.13), whereas controls for the remaining two *XRCC1* variants *XRCC1 Arg²⁸⁰ His* ($\chi^2=12.9$, **p=0.0003**) and *Gln⁶³² Gln* ($\chi^2=137.1$, p=0.00) showed deviation from HWE. The study did not show any significant difference in the genotypic frequency between cases and controls for the four (*XRCC1*) variants *i.e.* *Arg¹⁹⁴ Trp* ($\chi^2=2.66$, df=2, p=0.26), *Pro²⁰⁶ Pro* ($\chi^2=3.3$, df=2, p=0.19), *Arg²⁸⁰ His* ($\chi^2=137.1$, df=2, p=0.08), *Arg³⁹⁹ Gln* ($\chi^2=4.36$, df=2, p=0.113). In case of *XRCC1 Gln⁶³² Gln* there was a notable difference in the genotypic frequency between cases and controls ($\chi^2=36.2$, df=2, **p=<0.0001**).

In case of *XRCC1 Arg¹⁹⁴ Trp* there was over representation of the subjects with heterozygous (*Arg/Gln*) genotype (21.8%) as compared to lung cancer controls (16.9%). The minor allele frequency (MAF) was found to be higher in cases (0.12) as compared to controls (0.09). After applying logistic regression lung cancer risk analysis showed that subjects having heterozygous genotype (*C/T*) for *XRCC1 Arg¹⁹⁴ Trp* possessed a very moderate risk towards lung cancer though it was not statistically significant (OR=1.29, 95%C.I.=0.86-1.95, p=0.20). The frequency of subjects with homozygous variant (*Trp/Trp*) genotype was less represented in lung cancer cases therefore we combined both the heterozygotes (*Arg/Trp*) and homozygous (*Trp/Trp*) variants as a single group and compared it with the wild-type (*Arg/Arg*) genotype. This combination revealed no association towards lung cancer susceptibility (OR=1.25, 95%C.I.=0.84-1.87, p=0.26) as shown in table 5.10.

The next SNP that was evaluated was *XRCC1 Pro²⁰⁶ Pro* and it was found that subjects

carrying heterozygous genotype (*A/G*) were found to have a higher frequency in cases (49.6%) as compared to controls (42.7%), whereas subjects with mutant (*GG*) genotype were over-represented in controls (12.0%) as compared to cases (9.68%). The MAF was 0.35 and 0.33 in cases and controls respectively. Though there was a comparable difference between the frequency distribution of single allelic carriers (*Pro/Pro*) and mutant genotype (*Pro/Pro*) between cases and controls, the association analysis did not show any risk towards lung cancer when compared to the wild type genotype (*Pro/Pro*) as shown in table 5.10.

The next SNP evaluated in the present study was *XRCC1 Arg²⁸⁰His*, our data revealed that controls had a higher prevalence of the heterozygous (*GA*) genotype (82.4%) as compared to the lung cancer group (75.7%), whereas subjects homozygous for the variant A allele (*AA* genotype) were more common among cases (14.5%) as compared to controls (9.5%). The minor allele frequency was essentially similar in both cases and controls (0.52 vs 0.51). Taking the wild type (*GG*) genotype as a reference group no risk was found to be associated with heterozygous genotype (OR=0.82, 95%C.I.=0.47-1.45, p=0.51). Subjects carrying mutant genotype (*AA*) showed a trend towards increased risk when compared to wild type (*GG*) subjects though no statistical significance was observed (OR=1.45, 95%C.I.=0.70-3.02, p=0.31).

Furthermore, analysis revealed that for *XRCC1 Arg³⁹⁹Gln*, lung cancer patients had a higher frequency of representation for the wild type genotype (*GG*) as compared to controls (28.1% vs 24.3%), whereas subjects homozygous for the variant allele (*AA*) were more represented among controls (21.5%) as compared to cases (15.4%) as shown in table 5.10. Heterozygous (*GA*) subjects were found to be represented more in cases (56.3%) as compared to controls (54.1%). Minor allele frequency (MAF) was higher in controls (0.49) in comparison to cases (0.44). Logistic regression analysis indicated that subjects carrying the homozygous A allele (*AA*) were less likely to develop lung cancer, although the data did not fully reach statistical significance (OR=0.61, 95%C.I.=0.37-1.0, p=0.06). When both the heterozygous and mutant genotype were combined (*GA+AA*) a much reduced risk towards lung cancer (OR=0.25, 95%C.I.=0.56-1.16, p=0.25) was observed.

In case of polymorphic *XRCC1 Gln⁶³²Gln*, there was a higher representation of wild type (*GG*) subjects in controls (38.4%) as compared to cases (18.4%). Study subjects with both single allele (*GA*) (43.0% vs 39.0%) and double allelic variant (*AA*) (38.4 vs 22.7%) genotypes were found to be over-represented in cases as compared to controls respectively as shown in table 5.10. The minor allelic frequency of *XRCC1 Gln⁶³²Gln* was found to be higher in cases (0.60) as compared to controls (0.42). Association analysis revealed that subjects carrying single copy of variant allele (*GA*) had a 2-fold increased susceptibility

towards lung cancer which was found to be highly significant (OR=2.2, 95%C.I.=1.47-3.29, $p=0.0001$). Similarly, subjects with the variant mutant genotype (AA) were also found to increase the risk of lung cancer when compared to the homozygous wild-type genotype (GG) (OR=3.4, 95%C.I.=2.26-5.39, $p=<0.0001$). The combined variant genotype (OR=2.6, 95%C.I.=1.85-3.87, $p=<0.0001$) equally showed a high risk towards lung cancer.

Table 5.10: Polymorphic XRCC1 and Lung cancer susceptibility

Genes	Controls N=325 (n%)	Cases N=330 (n%)	Overall AOR (95% CI) ^a	p ^b
XRCC1¹⁹⁴ rs 1799782 C/T				
CC	267 (82.1)	256 (77.5)	1 (Reference)	
CT	55 (16.9)	72 (21.8)	1.29 (0.86-1.95)	0.20
TT	3 (0.92)	2 (0.60)	0.50 (0.075-3.38)	0.48
CT+TT	58 (17.8)	74 (22.4)	1.25 (0.84-1.87)	0.26
C	589	584		
G	61	76		
MAF	0.09	0.12		
XRCC1²⁰⁶ rs 915927 A/G				
AA	147 (45.2)	134 (40.6)	1(Reference)	
AG	139 (42.7)	164 (49.6)	12.0 (0.89-1.70)	0.18
GG	39 (12.0)	32 (9.68)	0.93 (0.54-1.61)	0.82
AG+GG	178 (54.7)	196 (59.3)	1.1 (0.86-1.65)	0.2
A	433	432		
G	217	228		
MAF	0.33	0.35		
XRCC1²⁸⁰ rs 25489 G/A				
GG	26 (8.0)	32 (9.6)	1 (Reference)	
GA	268 (82.4)	250 (75.7)	0.82 (0.47-1.45)	0.51
AA	31 (9.5)	48 (14.5)	1.45 (0.70-3.02)	0.31
GA+AA	299 (92)	298 (90.3)	0.89 (0.51-1.55)	0.68
G	320	314		
A	330	346		
MAF	0.51	0.52		
XRCC1³⁹⁹ rs 25487 G/A				
GG	79 (24.3)	93 (28.1)	1 (Reference)	
GA	176 (54.1)	186 (56.3)	0.88 (0.60-1.28)	0.50
AA	70 (21.5)	51 (15.4)	0.61 (0.37-1.00)	0.06
GA+AA	246 (75.6)	237 (71.7)	0.25 (0.56-1.16)	0.25
G	334	372		

to be cont'd on next page

Table 5.10: Polymorphic XRCC1 and Lung cancer susceptibility (contd..)

Genes	Controls N=325 (n%)	Cases N=330 (n%)	Overall AOR (95% CI) ^a	p ^b
A	316	288		
MAF	0.49	0.44		
XRCC1⁶³² rs 3547 G/A				
GG	124 (38.1)	61 (18.4)	1 (Reference)	
GA	127 (39.0)	142 (43.0)	2.2 (1.47-3.29)	0.0001
AA	74 (22.7)	127 (38.4)	3.4 (2.26-5.39)	< 0.0001
GA+AA	201 (61.8)	269 (81.5)	2.67 (1.85-3.87)	< 0.0001
G	375	264		
A	275	396		
MAF	0.42	0.60		

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)- Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.6.3 Genotypic and allelic frequency of *XRCC1* polymorphic gene and its association with lung cancer risk on the basis of histological subtypes

In order to segregate the study population into high risk subgroups, study was further comprehended on the basis of histology as shown in table 5.11, 5.12 and 5.13. In case of *XRCC1 Arg¹⁹⁴ Trp* it was observed that 74.5% of ADCC subjects had wild type (*CC*) genotype, 24.5% had heterozygous (*CT*) genotype and only 0.9% had double allelic variant (*TT*) genotype. On the same lines SQCC subtypes showed that there were 78.4% subjects with wild type (*CC*) genotype, 20.8% patients were of single allelic genotype (*CT*), and 0.7% had the mutant genotype (*TT*). Subjects with heterozygous (*CT*) genotype for both ADCC and SQCC subtype did not show any significant risk towards lung cancer as shown in tables 5.11 and 5.12. Association analysis showed that individuals carrying a single copy of variant allele (*CT*) for ADCC subtype showed a visible trend of increased risk (OR=1.43, 95% C.I.=0.82-2.5, p=0.20). However, SQCC subjects carrying mutant (*TT*) genotype were showing a protective effect for lung cancer (OR=0.42, 95% C.I.=0.039-4.67, p=0.48) though statistically non-significant as shown in table 5.12.

In case of *XRCC1 Pro²⁰⁶ Pro* the frequency distribution of wild type genotype (*AA*) for

ADCC patients stated there were 42.4% cases and 45.2% controls respectively. While there were 49.0% single allelic carrier (*AG*) subjects under cases and 42.7% subjects under controls. When odds ratio was calculated a trend of no association was observed for heterozygous subjects (*AG*) as shown in table 5.11. For SQCC subtype 42.4% subjects were wild type (*AA*) and 45.2% of the control subjects were also harboring the wild type alleles (*AA*). The percentage of subjects carrying heterozygous genotype (*AG*) for SQCC subtype was higher in cases (48.9%) as compared to controls (42.7%). Our data observed that only 8.63% of lung cancer subjects who were diagnosed with SQCC were carrying mutant genotype (*GG*) and (12.0%) control subjects harbored double allelic variant genotype (*GG*). However the difference was not significant enough to possess any risk towards lung cancer in case of SQCC histological subtype (OR=0.87, 95%C.I.=0.41-1.86, p=0.73). No significant association with any genotype was observed as odds were nearly 1 for both heterozygotes (*AG*) and subjects possessing double allele of variant genotype (*GG*) as indicated in table 5.12.

In case of SCLC histology it was evaluated that 51.8% SCLC patients were harboring the heterozygous genotype (*AG*) as compared to controls (42.7%). Taking wild type genotype (*AA*) as a reference it was observed that SCLC subjects carrying heterozygous genotype (*GA*) had a weak association with lung cancer (OR=1.72, 95%C.I.=0.97-3.04, p=0.06). Similarly, the mutant genotype (*GG*) did had a weak association towards lung cancer (OR=1.51, 95%C.I.=0.65-3.52, p=0.33). Furthermore, when both heterozygous and mutant genotype were combined as a single genotype (*AG+GG*) the lung cancer risk remained unchanged (OR=1.69, 95%C.I.=0.97-1.29, p=0.05) as shown in table 5.13.

In case of *XRCC1 Arg²⁸⁰His* frequency of subjects harboring heterozygous genotype (*GA*) and diagnosed with ADCC histology were higher in controls (82.4%) as compared to cases (75.4%). On the contrary, percentage of mutant genotype (*AA*) was found to be higher in ADCC subjects (16.9 %) as compared to controls (9.5%). The difference between the frequency of genotypes among cases and controls was not significant, therefore taking wild type genotype (*GG*) as a reference a visible trend of risk was observed (OR=2.33, 95%C.I.=0.80-6.70, p=0.11) as shown in table 5.11. In the case of SQCC subtype, subjects carrying both susceptible alleles (*AA*) exhibited a higher frequency of distribution as compared to controls (15.1% vs 9.5%). Multivariate regression analysis showed a visible trend of risk for SQCC subjects (OR=1.75, 95%C.I.=0.69-4.40, p=0.23) when compared with the wild type subjects (*GG*) as shown in table 5.12 .

In case of *XRCC1 Arg³⁹⁹Gln* a higher representation of heterozygous (*GA*) subjects was observed in controls (54.1%) as compared to ADCC (50.0%) subjects. Regression analysis indicated a protective effect in ADCC subtype for heterozygous (*GA*) genotype when

compared with the wild-type subjects (GG), but the analysis did not reach statistical significance (OR=0.61, 95%C.I.=0.355-1.07, $p=0.08$). The frequency of the mutant genotype (AA) was represented more in controls (21.5%) than in SQCC (12.2%) and SCLC (12.3%) patients. Association analysis revealed a visible trend of protective effect for both SQCC (OR=0.50, 95%C.I.=0.24-1.02, $p=0.05$) and SCLC (OR=0.47, 95%C.I.=0.19-1.50, $p=0.08$) subtype of lung cancer.

In case of *XRCC1 Gln⁶³²Gln* our study reported that ADCC subjects had a higher representation in all the three genotypes (GA , AA , $GA+AA$) in cases as compared to controls. 40.5% of ADCC subjects had single susceptible allele (GA), 38.6% subjects had both variant allele (AA) and 79.2% subjects carried combined variant genotype ($GA+AA$). Controls showed 39.0% subjects were single allelic carriers (GA), 22.7% were double allelic variant carriers (AA) and 61.8% were combined variant carriers ($GA+AA$). Lung cancer cases diagnosed with ADCC and carrying single copy of variant allele (GA) exhibited a two fold risk (OR=1.85, 95%C.I.=1.03-3.32, **$p=0.039$**). The risk further amplified to 3-fold for subjects with both allele of variant genotype (AA) (OR=3.09, 95%C.I.=1.03-3.32, **$p=0.0003$**).

For SQCC histology, it was observed that 16.5% patients had wild type genotype (GG) genotype and 42.2% had heterozygous genotype (GA) and 41.0% patients carried mutant genotype (AA). On the contrary 39.0% control subjects carried single susceptible allele (GA), 22.7% subjects were double allelic variant carriers (AA) and 61.8% were combined variant carriers ($GA+AA$). The difference in the frequency distribution for SQCC patients with single susceptible allele (GA), mutant genotype (AA) and combined variant genotype ($GA+AA$) was significantly different between cases and controls as shown in table 5.12. Therefore, heterozygous subjects acquired a significant risk (OR=2.54, 95%C.I.=1.45-4.46, **$p=0.001$**) towards lung cancer. Further, the risk elevated to four folds for mutant genotype (AA) (OR=4.16, 95%C.I.=2.30-7.52, **$p=<0.001$**) when wild (GG) genotype is used as a reference as shown in table 5.12. Similar to the two histology, subjects with SCLC histology and carrying heterozygous genotype (GA) exhibited a higher frequency compared to controls (49.3 vs 39.0%) therefore a high risk was observed (OR=2.51, 95%C.I.=1.28-4.90, **$p=0.0069$**). On the same lines there was a high representation of mutant genotype (AA) compared to controls (30.8% vs 22.7%) exhibiting a high risk when compared to controls (OR=2.45, 95%C.I.=1.16-5.90, **$p=0.0182$**) as shown in table 5.13.

Table 5.11: Genotypic distribution of *XRCC1* genetic variants and their association with risk of lung cancer based on ADCC histological subtype

Genes		ADCC		
XRCC1¹⁹⁴ C/T	Controls N=325 (n%)	Cases N=106 (n%)	AOR (95% CI)^a	p^b
CC	267 (82.1)	79 (74.5)	1 (Reference)	
CT	55 (16.9)	26 (24.5)	1.43 (0.82-2.5)	0.20
TT	3 (0.92)	1 (0.90)	0.92 (0.080-10.73)	0.95
CT+TT	58 (17.8)	27 (25.4)	1.41 (0.81-2.6)	0.21
XRCC1²⁰⁶ A/G	Controls N=325 (n%)	Cases N=106 (n%)	AOR (95% CI)^a	p^b
AA	147 (45.2)	45 (42.4)	1 (Reference)	
AG	139 (42.7)	52 (49.0)	1.18 (0.73-1.19)	0.47
GG	39 (12.0)	9 (8.49)	0.77 (0.34-1.72)	0.55
AG+GG	178 (54.7)	61 (57.5)	1.12 (0.70-1.77)	0.62
XRCC1²⁸⁰ G/A	Controls N=325 (n%)	Cases N=106 (n%)	AOR (95% CI)^a	p^b
GG	26 (8.0)	8 (7.5)	1 (Reference)	
GA	268 (82.4)	80 (75.4)	1.09 (0.46-2.60)	0.83
AA	31 (9.5)	18 (16.9)	2.33 (0.80-6.70)	0.11
GA+AA	299 (92)	98 (92.4)	1.20 (0.51-2.83)	0.66
XRCC1³⁹⁹ G/A	Controls N=325 (n%)	Cases N=106 (n%)	AOR (95% CI)^a	p^b
GG	79 (24.3)	31 (29.2)	1 (Reference)	
GA	176 (54.1)	53 (50.0)	0.61 (0.35-1.70)	0.08
AA	70 (21.5)	22 (20.7)	0.74 (0.38-1.44)	0.38
GA+AA	246 (75.6)	75 (70.7)	0.66 (0.39-1.11)	0.11
XRCC1⁶³² G/A	Controls N=325 (n%)	Cases N=106 (n%)	Overall AOR (95% CI)^a	p^b
GG	124 (38.1)	22 (20.7)	1 (Reference)	
GA	127 (39.0)	43 (40.5)	1.85 (1.03-3.32)	0.039
AA	74 (22.7)	41 (38.6)	3.09 (1.68-5.69)	0.0003
GA+AA	201 (61.8)	84 (79.2)	2.31 (1.35-3.94)	0.0021

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

Table 5.12: Genotypic distribution of *XRCC1* genetic variants and their association with risk of lung cancer based on SQCC histological subtype

Genes		SQCC		
XRCC1¹⁹⁴ C/T	Controls N=325 (n%)	Cases N=139 (n%)	AOR (95% CI)^a	p^b
CC	267 (82.1)	109 (78.4)	1 (Reference)	
CT	55 (16.9)	29 (20.8)	1.23 (0.72-2.10)	0.43
TT	3 (0.92)	1 (0.71)	0.42 (0.039-4.67)	0.48
CT+TT	58 (17.8)	30 (21.5)	1.18 (0.70-1.99)	0.53
XRCC1²⁰⁶ A/G	Controls N=325 (n%)	Cases N=139 (n%)	AOR (95% CI)^a	p^b
AA	147 (45.2)	59 (42.4)	1 (Reference)	
AG	139 (42.7)	68 (48.9)	1.21 (0.78-1.89)	0.37
GG	39 (12.0)	12 (8.63)	0.87 (0.41-1.86)	0.73
AG+GG	178 (54.7)	80 (57.5)	1.15 (0.75-1.76)	0.50
XRCC1²⁸⁰ G/A	Controls N=325 (n%)	Cases N=139 (n%)	AOR (95% CI)^a	p^b
GG	26 (8.0)	14 (10.0)	1 (Reference)	
GA	268 (82.4)	104 (74.8)	0.85 (0.41-1.75)	0.67
AA	31 (9.5)	21 (15.1)	1.75 (0.69-4.40)	0.23
GA+AA	299 (92)	125 (89.9)	0.93 (0.46-1.91)	0.86
XRCC1³⁹⁹ G/A	Controls N=325 (n%)	Cases N=139 (n%)	AOR (95% CI)^a	p^b
GG	79 (24.3)	38 (27.3)	1 (Reference)	
GA	176 (54.1)	84 (60.4)	1.03 (0.63-1.69)	0.88
AA	70 (21.5)	17 (12.2)	0.50 (0.24-1.02)	0.05
GA+AA	246 (75.6)	101 (72.6)	0.89 (0.55-1.43)	0.63
XRCC1⁶³² G/A	Controls N=325 (n%)	Cases N=139 (n%)	AOR (95% CI)^a	p^b
GG	124 (38.1)	23 (16.5)	1 (Reference)	
GA	127 (39.0)	59 (42.4)	2.54 (1.45-4.46)	0.001
AA	74 (22.7)	57 (41.0)	4.16 (2.30-7.52)	< 0.0001
GA+AA	201 (61.8)	116 (83.4)	3.14 (1.86-5.28)	< 0.0001

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

Table 5.13: Genotypic distribution of *XRCC1* genetic variants and their association with risk of lung cancer based on SCLC histological subtype

Genes		SCLC		
XRCC1¹⁹⁴ C/T	Controls N=325 (n%)	Cases N=81 (n%)	AOR (95% CI)^a	p^b
CC	267 (82.1)	66 (81.4)	1 (Reference)	
CT	55 (16.9)	15 (18.5)	0.91 (0.46-1.81)	0.80
TT	3 (0.92)			
CT+TT	58 (17.8)	15 (8.5)	0.85 (0.43-1.68)	0.64
XRCC1²⁰⁶ A/G	Controls N=325 (n%)	Cases N=81 (n%)	AOR (95% CI)^a	p^b
AA	147 (45.2)	28 (34.5)	1 (Reference)	
AG	139 (42.7)	42 (51.8)	1.72 (0.97-3.04)	0.06
GG	39 (12.0)	11 (13.5)	1.51 (0.65-3.52)	0.33
AG+GG	178 (54.7)	53 (65.3)	1.69 (0.97-2.90)	0.05
XRCC1²⁸⁰ G/A	Controls N=325 (n%)	Cases N=81 (n%)	AOR (95% CI)^a	p^b
GG	26 (8.0)	10 (12.3)	1 (Reference)	
GA	268 (82.4)	63 (77.7)	0.64 (0.28-1.45)	0.29
AA	31 (9.5)	8 (9.87)	0.79 (0.23-2.67)	0.71
GA+AA	299 (92)	71 (87.6)	0.66 (0.29-1.48)	0.31
XRCC1³⁹⁹ G/A	Controls N=325 (n%)	Cases N=81 (n%)	AOR (95% CI)^a	p^b
GG	79 (24.3)	24 (29.6)	1 (Reference)	
GA	176 (54.1)	47 (58.2)	0.82 (0.44-1.50)	0.53
AA	70 (21.5)	10 (12.3)	0.47 (0.19-1.50)	0.08
GA+AA	246 (75.6)	57 (70.3)	0.73 (0.49-1.32)	0.30
XRCC1⁶³² G/A	Controls N=325 (n%)	Cases N=81 (n%)	AOR (95% CI)^a	p^b
GG	124 (38.1)	16 (19.7)	1 (Reference)	
GA	127 (39.0)	40 (49.3)	2.51 (1.28-4.90)	0.0069
AA	74 (22.7)	25 (30.8)	2.45 (1.16-5.90)	0.0182
GA+AA	201 (61.8)	65 (80.2)	1.50 (1.32-4.61)	0.0043

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.6.4 Effect of smoking on polymorphic *XRCC1* variants and lung cancer risk associated

In order to study the gene environment interaction, study subjects were divided into smokers and non-smokers on the basis of their smoking status. The genotypic frequency distribution in relation to smoking for the five SNPs is summarized in table 5.14 and 5.15. In case of *XRCC1 Arg¹⁹⁴Trp* the frequency of lung cancer subjects who were smokers and carrying the heterozygous genotype (*CT*) for cases was 21.6% as compared to 17.0% for controls. Taking wild type genotype as a reference no risk was observed towards lung cancer for single allelic carriers (*CT*) (OR=1.27, 95%C.I.=0.80-2.0, p=0.29) among smokers. However, lung cancer diseased subjects harboring mutant genotype (*TT*) for the above SNP had a lesser frequency compared to controls (0.7% vs 1.3%). Further, a protective trend was observed for smokers who were carrying the mutant (*TT*) genotype (OR=0.53, 95%C.I.=0.082-3.5, p=0.51) as compared to the wild type genotype (*CC*).

In case of *XRCC1 Pro²⁰⁶Pro* it was observed, that patients carrying heterozygous genotype (*AG*) had a higher representation of heterozygous genotype in comparison to controls among both the smokers (50.7% vs 43.0%) and non-smokers (45.1% vs 42.1%) subgroup. On the contrary as shown in table 5.14 decreased frequency of mutant genotype (*GG*) was observed among smokers (8.9%) when compared to controls (11.2%) but the association analysis exhibited a lack of any significant correlation between smoking and the polymorphic site in correlation to lung cancer risk. However, non-smokers carrying mutant genotype (*GG*) had a lesser frequency of representation with respect to control subjects (12.9% vs 13.7%) and therefore, a strong protective effect was observed (OR=0.38, 95%C.I.=0.15-0.96, **p=0.04**) as shown in table 5.15.

In case of *XRCC1 Arg²⁸⁰His* neither smokers nor non-smokers showed any significant association towards lung cancer risk as shown in table 5.14, and 5.15. Within the smokers group, it was observed that lung cancer subjects carrying heterozygous (*GA*) genotype were represented more in controls as compared to cases (83.4 vs 75.7%). However, the difference was not significant and therefore no association was observed. The percentage of patient subjects with single allele of variant genotype (*GA*) was 75.8% which was found to be lower as compared to controls (80.3%). Henceforth, a trend of protective effect was observed in case of non-smokers with heterozygous genotype (*GA*) (OR=0.47, 95%C.I.=0.13-1.65, p=0.24). Furthermore combined variants *i.e.* heterozygous and mutant genotype (*GA+AA*) among smokers were found to be equally distributed in both cases (90.6%) and controls (90.5%) respectively. Therefore no significant association was

observed among smokers as shown in table 5.14. In non-smokers, controls had a higher frequency (95.0%) of both the combined heterozygous and mutant genotype ($GA+AA$) as compared to cases (88.7%). Therefore, similar to heterozygous subjects a trend towards a protective effect was observed (OR=0.46, 95%C.I.=0.13-1.63, $p=0.23$).

XRCC1 Arg³⁹⁹Gln polymorphism also did not affect lung cancer susceptibility in both smokers and non-smokers. Lung cancer subjects who were smokers and carried double allelic variant (AA) genotype were found to be less represented in controls (15.2%) as compared to cases (20.1%). Logistic regression analysis showed no association or correlation towards lung cancer risk and smoking as shown in table 5.14. Non-smokers with mutant genotype had a comparable difference in the frequency distribution between cases (16.1%) and controls (24.5%). Taking wild type genotype as a reference a visible trend of protective effect was observed (OR=0.44, 95%C.I.=0.14-1.34, $p=0.15$).

In case of *XRCC1 Gln⁶³²Gln* polymorphism it was observed that smokers for all the three genotypic combinations unanimously showed an increased risk towards lung cancer as shown in table 5.14. Mutant (GG) subjects who were smokers were found to be more represented as compared to control subjects who were smokers too (38.0% vs 22.8%). Lung cancer subjects who were smokers and harboring the mutant genotype (GG) were found to have four times high risk of developing lung cancer (OR=3.7, 95%C.I.=2.2-6.1, $p<0.0001$) as compared to control subjects who were smokers too. Double allelic variants (GG) within non-smokers showed a 3 fold risk as there was a high representation of mutant subjects in cases (40.3%) as compared to controls (22.5%) (OR=2.75, 95%C.I.=1.1-6.8, $p=0.02$) as depicted in table 5.15.

Table 5.14: Genotypic distribution of *XRCC1* variants based on smoking status and its association with risk of Lung cancer (Smokers)

Genes	Smokers			
	Controls N=223 (n%)	Cases N=268 (n%)	AOR (95% CI) ^a	p ^b
XRCC1¹⁹⁴ C/T				
CC	182 (81.6)	208 (77.6)	1.0	
CT	38 (17.0)	58 (21.6)	1.27 (0.80-2.0)	0.29
TT	3 (1.3)	2 (0.7)	0.53 (0.08-3.5)	0.51
CT+TT	41 (18.3)	60 (22.3)	1.22 (0.77-1.93)	0.37
XRCC1²⁰⁶ A/G				
AA	102 (45.7)	108 (40.3)	1.0	
AG	96 (43.0)	136 (50.7)	1.34 (0.91-1.97)	0.12

to be cont'd on next page

Table 5.14: Genotypic distribution of *XRCC1* variants based on smoking status and its association with risk of Lung cancer (Smokers) (contd..)

Genes	Smokers			
GG	25 (11.2)	24 (8.9)	0.89 (0.47-1.69)	0.73
AG+GG	121 (54.2)	160 (59.7)	1.25 (0.86-1.81)	0.22
XRCC1²⁸⁰ G/A	Controls N=223 (n%)	Cases N=268 (n%)	AOR (95% CI)^a	p^b
GG	21 (9.41)	25 (9.32)	1.0	
GA	186 (83.4)	203 (75.7)	0.92 (0.49-1.7)	0.81
AA	16 (17.1)	40 (14.9)	1.9 (0.82-4.43)	0.12
GA+AA	202 (90.5)	243 (90.6)	1.00 (0.54-1.87)	0.97
XRCC1³⁹⁹ G/A	Controls N=223 (n%)	Cases N=268 (n%)	AOR (95% CI)^a	p^b
GG	59 (26.4)	77 (28.7)	1.0	
GA	119 (53.3)	150 (55.9)	0.92 (0.60-1.41)	0.72
AA	45 (20.1)	41 (15.2)	0.68 (0.38-1.1)	0.18
GA+AA	164 (73.5)	191 (71.2)	0.86 (0.57-1.29)	0.47
XRCC1⁶³² G/A	Controls N=223 (n%)	Cases N=268 (n%)	AOR (95% CI)^a	p^b
GG	85 (38.1)	46 (17.1)	1.0	
GA	87 (39.0)	120 (44.7)	2.6 (1.65-4.20)	0.0001
AA	51 (22.8)	102 (38.0)	3.7 (2.2-6.1)	<0.0001
GA+AA	138 (61.8)	222 (82.8)	3.08(1.96-4.63)	<0.0001

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

Table 5.15: Genotypic distribution of *XRCC1* variants based on smoking status and its association with risk of Lung cancer (Non-smokers)

Genes	Non Smokers			
XRCC1¹⁹⁴ C/T	Controls N=102 (n%)	Cases N=62 (n%)	AOR (95% CI)^a	p^b
CC	85 (83.3)	48 (77.4)	1	
CT	17 (16.6)	14 (22.5)	1.61 (0.64-4.00)	0.30
TT	-	-	-	
CT+TT	17 (16.6)	14 (22.5)	1.61 (0.64-4.00)	0.30

to be cont'd on next page

Table 5.15: Genotypic distribution of *XRCC1* variants based on smoking status and its association with risk of Lung cancer (Non-smokers) (contd..)

Genes	Non Smokers			
XRCC1²⁰⁶ A/G	Controls N=102 (n%)	Cases N=62 (n%)	AOR (95% CI)^a	p^b
AA	45 (44.1)	26 (41.9)	1	
AG	43 (42.1)	28 (45.1)	0.57 (0.27-1.17)	0.127
GG	14 (13.7)	8 (12.9)	0.38 (0.15-0.96)	0.04
AG+GG	57 (55.8)	36 (58.0)	0.67 (0.35-1.31)	0.25
XRCC1²⁸⁰ G/A	Controls N=102 (n%)	Cases N=62 (n%)	AOR (95% CI)^a	p^b
GG	5 (4.9)	7 (11.2)	1	
GA	82 (80.3)	47 (75.8)	0.47 (0.13-1.65)	0.24
AA	15 (14.7)	8 (12.9)	0.46 (0.093-2.27)	0.34
GA+AA	97 (95.0)	55 (88.7)	0.46 (0.13-1.63)	0.23
XRCC1³⁹⁹ G/A	Controls N=102 (n%)	Cases N=62 (n%)	AOR (95% CI)^a	p^b
GG	20 (19.6)	16 (25.8)	1	
GA	57 (55.8)	36 (58.0)	0.59 (0.59-1.41)	0.24
AA	25 (24.5)	10 (16.1)	0.44 (0.14-1.34)	0.15
GA+AA	82 (80.3)	46 (74.1)	0.55 (0.24-1.20)	0.158
XRCC1⁶³² G/A	Controls N=102 (n%)	Cases N=62 (n%)	AOR (95% CI)^a	p^b
GG	39 (38.2)	15 (24.1)	1	
GA	40 (39.2)	22 (35.4)	1.39 (0.58-3.1)	0.48
AA	23 (22.5)	25 (40.3)	2.76 (1.1-6.8)	0.02
GA+AA	63 (61.7)	47 (75.8)	1.80 (0.87-3.9)	0.10

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

5.6.5 Relationship of genetic variants in *XRCC1* gene with clinicopathological parameters

Tables 5.16 and 5.17 below show the correlation of five polymorphic variants of *XRCC1* and various clinicopathological features such as tumor stage, primary tumor extension, lymph node invasion and metastasis. Various polymorphic sites of *XRCC1* were analyzed for the advancement of these clinical parameters and the risk was evaluated using logistic regression analysis. Wild type genotype was used as a reference and odds for the advancement of the disease was calculated. The odds obtained was adjusted for factors like age, gender and smoking. The diseased study patients were stratified into clinical stage III and stage IV. For primary tumor extension lung cancer subjects were classified into TX+T1+T2 and T3+T4 groups. For lymph nodes involved study subjects were divided into N0 and N1+N2+N3 stages of lymph node involvement as the other group where cancer has spread to lymph nodes. On the basis of metastasis the study was classified into M0 and M1 wherein tumor has not spread to other parts of body in case of M0 whereas M1 group had tumor spread to other parts of lung and body.

As shown in table 5.16 there was no correlation of five polymorphic sites with the stage of the tumor. In case of *XRCC1 Arg¹⁹⁴ Trp* the frequency distribution for subjects with single allelic variant genotype (*CT*) was nearly similar for both stage III (20.8%) and stage IV (21.5%) patients. Distribution of mutant genotype (*TT*) was found to be similar in both stage III (0.5%) and stage IV (0.6%) subjects. Regression analysis did not show any significant association for stage and tumor for both heterozygous and mutant subjects as shown in table 5.16.

XRCC1 Pro²⁰⁶ Pro polymorphism was analyzed for its correlation with various clinicopathological factors. The study subjects carrying single allelic variant genotype (*AG*) and diagnosed with clinical stage (IV) had a higher representation (50.6%) as compared to subjects with stage (III) (47.9%) of the disease. Since the difference was not significant logistic regression analysis suggested no association towards lung cancer (OR=1.14, 95% C.I.=0.70-1.85, p=0.58) as shown in table 5.16.

In case of *XRCC1 Arg²⁸⁰ His* frequency of single allelic variant (*GA*) genotype was 77.7% and 74.5% among stage III and stage IV respectively. The association analysis suggested no correlation with lung cancer risk as shown in table 5.16. In case of *XRCC1 Arg³⁹⁹ Gln* the percentage of subjects carrying heterozygous (*GA*) genotype was 59.5% for stage III subjects and 52.0% for stage IV subjects. Regression analysis indicated no involvement of stage and the polymorphism towards lung cancer predisposition as shown in table 5.16. In case of polymorphic variant *XRCC1 Gln⁶³² Gln* patients with heterozygous genotype

(*GA*) and stage IV had a slightly higher frequency (45.1%) as compared to stage III patients (42.1%). Logistic regression analysis revealed, a trend towards moderate risk (OR=1.38, 95%C.I.=0.72-2.62, $p=0.31$).

Primary tumor extension was the next parameter evaluated for the five polymorphic sites as shown in table 5.17. The investigative study of *XRCC1 Pro*²⁰⁶*Pro* for its correlation with primary tumor extension, suggested that the frequency of subjects harboring single allele of variant genotype (*GA*) was consistent among both the two groups named TX+T1+T2 (59.4%) and T3+T4 (59.5%) respectively. Therefore, no significant association was observed between the polymorphic variant and primary tumor extension (OR=1.14, 95%C.I.=0.70-1.85, $p=0.58$). In case of *XRCC1 Arg*²⁸⁰*His* subjects diagnosed with T3+T4 tumor size and carrying mutant genotype (*GG*) had a higher frequency (15.7%) as compared to subjects with TX+T1+T2 tumor size (10.1%). When compared to the wild type genotype (*GG*) no correlation towards lung cancer was observed (OR=1.46, 95%C.I.=0.79-2.7, $p=0.22$). For polymorphic variant, *XRCC1 Arg*³⁹⁹*Gln* the representation of both the heterozygous and mutant genotype as a single group (*GA+AA*) was found to be similar for both TX+T1+T2 (71.0%) and T3+T4 (72.7%) groups. Therefore no significant relationship between variant genotype and lymph node involvement was observed. On the other hand, combined variants (*GA+AA*) for polymorphic site *XRCC1 Gln*⁶³²*Gln* had comparable difference in the frequency for both TX+T1+T2 (84.0%) and T3+T4 (79.3%) groups, however the difference was not significant (OR=0.68, 95%C.I.=0.33-1.42, $p=0.31$).

Table 5.17 depicts the genotypic frequencies of polymorphic *XRCC1* in relation to parameters such as lymph node involved and metastasis. Some of the polymorphic variant of *XRCC1* summarized in table 5.17 exhibited involvement in lymph node invasion. In case of polymorphic variant *XRCC1 Pro*²⁰⁶*Pro* lung cancer patients who had lymph node involvement (N1+N2+N3) and carried the heterozygous genotype (*AG*), were found to be more represented than those lung cancer subjects with no lymph node involvement (51.7% vs 38.4%). Regression analysis showcased a positive correlation of lymph node involvement with subjects harboring single allelic variant (*AG*) genotype (OR=1.79, 95%C.I.=0.093-3.45, $p=0.07$) with nearly approaching significance. It was also observed that lung cancer subjects who were carriers of double allelic variant genotype (*AA*) for *XRCC1 Arg*²⁸⁰*His* and with lymph node involvement were found to be represented more as compared to subjects within N0 group (15.8% vs 7.6%). Considering the wild type genotype (*GG*) as a reference, subjects with mutant genotype (*AA*) were positively correlated with lymph node invasion (OR=2.16, 95%C.I.=1.07-4.31, **$p=0.02$**). The risk amplified when both mutant (*AA*) and heterozygous (*GA*) genotype were combined (*GA+AA*) as

a single genotype (OR=2.3, 95%C.I.=0.97-3.7, p=0.06) as indicated in table 5.17.

In case of *XRCC1 Arg³⁹⁹Gln* patients carrying mutant genotype (AA) were lesser represented for N1+N2+N3 (13.5%) category compared to subjects with N0 lymph node (23.0%). Therefore, for *XRCC1 Arg³⁹⁹Gln* subjects having N1+N2+N3 lymph nodes involvement exhibited no correlation with lymph node infiltration (OR=0.53, 95%C.I.=0.33-0.85, ***p*=0.009**).

The relation between metastatic status and polymorphic variants is detailed in table 5.17. The genotypic frequencies of variant genotype for *XRCC1 Arg¹⁹⁴Trp* exhibited no significant difference between the patients who showed metastasis (22.5%) and those showing absence of metastasis (20.7%). Hence, this polymorphic site did not show any association with spread of lung cancer to other organs. Other *XRCC1* variants evaluated in the study did not show any correlation with metastasis as shown in table 5.17. In case of *XRCC1 Pro²⁰⁶Pro* lung cancer subjects with distant metastasis and carrying mutant genotype were found to have similar frequency as compared to subjects without distant metastasis (9.0 vs 10.6%). Similarly, in case of *XRCC1 Arg²⁸⁰His* the percentage of subjects having combined variant genotype *GA+AA* genotype were found to be similar in both metastatic (91.7%) and non-metastatic (91.5%) subjects. Further, in case of *XRCC1 Gln⁶³²Gln*, the proportion of mutant subjects were similar in both metastatic and non-metastatic group (36.8% vs 37.0%). The frequency distribution for combined variant *GA+AA* was found not to be different as almost similar frequency was observed for the two metastatic groups (81.2% vs 79.7%). Hence, none of the polymorphic site contributed towards the progression of lung cancer among other organs as shown in table 5.17.

Table 5.16: Relation between clinicopathological factors and polymorphic *XRCC1* (Stage and primary tumor extension)

Genotype	Clinical stage		AOR (95% CI) ^a	p ^b	Primary tumor extension		AOR (95% CI) ^a	p ^b
	III	IV			TX+ T1+ T2	T3+ T4		
XRCC1 ¹⁹⁴ C/T	173	144			69	242		
CC	136 (78.6)	112 (77.7)	1		56 (81.1)	188 (77.6)	1	
CT	36 (20.8)	31 (21.5)	1.12 (0.64-1.95)	0.68	12 (17.3)	53(21.9)	1.30 (0.64-2.6)	0.45
TT	1 (0.5)	1 (0.6)	1.25 (0.31-5.0)	0.75	1 (1.4)	1 (0.4)	0.49 (0.12-2.0)	0.32
CT+TT	37 (21.3)	32 (22.2)	1.13 (0.67-1.9)	0.63	13 (18.8)	54 (22.3)	1.12 (0.58-2.13)	0.72
XRCC1 ²⁰⁶ A/G	173	144			69	242		
AA	71 (41.0)	58 (40.2)	1		28 (40.5)	98 (40.4)	1	
AG	83 (47.9)	73 (50.6)	1.14 (0.70-1.85)	0.58	35 (50.7)	119 (49.1)	1.10 (0.56-1.77)	0.93
GG	19 (10.9)	13 (9.0)	0.84 (0.55-1.28)	0.43	6 (8.6)	25 (10.3)	1.10 (0.67-1.81)	0.69
AG+GG	102 (58.9)	86 (59.7)	1.06 (0.66-1.69)	0.79	41 (59.4)	144 (59.5)	1.0 (0.58-1.75)	0.96
XRCC1 ²⁸⁰ G/A	173	144			69	242		
GG	15 (8.6)	15 (10.4)	1		7 (10.1)	19 (7.8)	1	
GA	129 (74.5)	112 (77.7)	0.86 (0.39-1.90)	0.72	55 (79.7)	185 (76.4)	1.20 (0.47-3.04)	0.68
AA	29 (16.7)	17 (11.8)	0.73 (0.45-1.20)	0.22	7(10.1)	38 (15.7)	1.46 (0.79-2.7)	0.22
GA+AA	159 (91.9)	129 (89.5)	0.33 (0.18-0.61)	0.0004	62 (89.8)	223 (92.1)	1.28 (0.51-3.21)	0.59
XRCC1 ³⁹⁹ G/A	173	144			69	242		
GG	48 (27.7)	41 (28.4)	1		20 (28.9)	66 (27.2)	1	
GA	103 (59.5)	75(52.0)	0.82 (0.48-1.41)	0.47	35 (50.7)	141 (58.2)	0.87 (0.58-1.31)	0.53
AA	22 (12.7)	28 (17.3)	1.20 (0.84-1.71)	0.31	14 (20.2)	35 (14.4)	0.87 (0.58-1.31)	0.53
GA+AA	125 (72.2)	103 (71.5)	0.93 (0.56-1.55)	0.80	49 (71.0)	176 (72.7)	1.09 (0.6-1.98)	0.7
XRCC1 ⁶³² G/A	173	144			69	242		
GG	35 (20.2)	26 (18.0)	1		11(15.9)	50(20.6)	1	
GA	73 (42.1)	65 (45.1)	1.38 (0.72-2.62)	0.31	34 (49.2)	101 (41.7)	0.61 (0.28-1.33)	0.21
AA	65 (37.5)	53 (36.8)	1.19 (0.81-1.56)	0.46	24 (34.7)	91 (37.6)	0.90 (0.60-1.35)	0.62
GA+AA	138 (79.7)	118 (81.9)	1.08 (0.79-1.48)	0.62	58 (84.0)	192 (79.3)	0.68 (0.33-1.42)	0.31

to be cont'd on next page

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

Table 5.17: Relation between clinicopathological factors and polymorphic *XRCC1* (Lymph node and metastasis)

Genotype	Lymph Node		AOR (95% C.I.) ^a	p ^b	Metastasis		AOR (95% C.I.) ^a	p ^b
	N0	N1+ N2+ N3			M0	M1		
XRCC1 ¹⁹⁴ C/T		259			178	133		
CC	42 (80.7)	202 (77.9)	1		141 (79.2)	103 (77.4)	1	
CT	8 (15.3)	57 (22.0)	1.30 (0.64-2.61)	0.45	36 (20.2)	29 (21.8)	1.18 (0.66-2.0)	0.56
TT	2 (3.8)	-	-	-	1 (0.5)	1 (0.7)	1.31 (0.32-5.30)	0.7
CT+TT	10 (19.2)	57 (22.0)	1.21 (0.61-2.4)	0.56	37 (20.7)	30 (22.5)	1.19 (0.70-2.03)	0.51
XRCC1 ²⁰⁶ A/G	52	259			178	133		
AA	26 (50.0)	100 (42.8)	1		70 (39.3)	56 (42.1)	1	
AG	20 (38.4)	134 (51.7)	1.79 (0.93-3.45)	0.07	89 (50.0)	65 (48.8)	0.96 (0.58-1.58)	0.88
GG	6 (11.5)	25 (9.6)	1.05 (0.63-1.73)	0.84	19 (10.6)	12 (9.0)	0.84 (0.55-1.28)	0.84
AG+GG	26 (50.0)	159 (61.3)	1.60 (0.87-2.93)	0.12	108 (60.6)	77 (57.8)	0.91 (0.57-1.47)	0.72
XRCC1 ²⁸⁰ G/A	52	259			178	133		
GG	8 (15.3)	18 (6.9)	1		15 (8.4)	11(8.2)	1	
GA	40 (76.9)	200 (77.2)	2.1 (0.85-5.29)	0.10	133 (74.7)	107 (80.4)	1.16 (0.49-2.7)	0.72
AA	4 (7.6)	41 (15.8)	2.16 (1.07-4.31)	0.02	30 (16.8)	15 (11.2)	0.83 (0.49-1.4)	0.49
GA+AA	44 (84.6)	241 (93.0)	2.3 (0.97-3.7)	0.06	163 (91.5)	122 (91.7)	1.09 (0.46-2.54)	0.83
XRCC1 ³⁹⁹ G/A	52	259			178	133		
GG	9 (17.3)	77 (29.7)	1		47 (26.4)	39 (29.3)	1	
GA	29 (55.7)	147 (56.7)	0.62 (0.28-1.41)	0.26	104 (58.4)	72 (54.1)	0.79 (0.45-1.36)	0.39
AA	12 (23.0)	35 (13.5)	0.53 (0.33-0.85)	0.009	27 (15.1)	22 (16.5)	0.95 (0.66-1.37)	0.81
GA+AA	41 (78.8)	182 (70.2)	0.5 (0.23-1.09)	0.08	131 (73.5)	94 (70.6)	0.82 (0.48-1.37)	0.45

to be cont'd on next page

Table 5.17: Relation between clinicopathological factors and polymorphic *XRCC1* (Lymph node and metastasis)

Genotype	Lymph node		AOR (95% CI) ^a	p ^b	Metastasis		AOR (95% CI) ^a	p ^b
	N0	N1+ N2+ N3			M0	M1		
XRCC1 ⁶³² G/A	52	259			178	133		
GG	9 (17.3)	52 (20.0)	1		36 (20.2)	25 (18.7)	1	
GA	21 (40.3)	114 (44.0)	0.85 (0.36-2.03)	0.72	76 (42.6)	59 (44.3)	1.31 (0.68-2.51)	0.41
AA	22 (42.3)	93 (35.9)	0.83 (0.53-1.29)	0.41	66 (37.0)	49 (36.8)	1.10 (0.79-1.53)	0.56
GA+AA	43 (82.6)	207 (79.9)	0.82 (0.54-1.24)	0.35	142 (79.7)	108 (81.2)	1.26 (0.69-2.29)	0.44

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

5.6.6 Combinatorial analysis of risk associated with different polymorphic sites of *XRCC1* gene

We further evaluated the effect of multiple combinations *i.e.* double and triple combinations of the five SNPs of *XRCC1* *i.e.* *Arg*¹⁹⁴*Trp*, *Pro*²⁰⁶*Pro*, *Arg*²⁸⁰*His*, *XRCC1 Arg*³⁹⁹*Gln* and *His*⁶³²*His* towards modulating lung cancer susceptibility as shown in table 5.18. Interactions were assessed for all doublet combinations but few of them were significantly modulating and elevating the lung cancer risk whereas rest were giving a visualized trend of either no effect or both having protective effect.

In case of genotypic combination of *XRCC1 Arg*³⁹⁹*Gln* and *XRCC1 Arg*¹⁹⁴*Trp* no risk was observed towards lung cancer as there was no significant difference in the distribution frequency between cases and controls for any genotypic combination as shown in table 5.18. In case of combination *XRCC1 Arg*³⁹⁹*Gln* and *Pro*²⁰⁶*Pro* the frequency of subjects carrying both variant alleles (*AA/GG*) between cases and controls for the above genotypic combination showed a higher frequency for cases (3.2%) as compared to controls (1.8%). Logistic regression analysis suggested a trend of high risk for lung cancer when compared to subjects carrying wild type (*GG/AA*) genotype (OR=2.17, 95%C.I.=0.31-15.14, p=0.43). When mutant (*AA/GG*) genotypes for both SNPs were combined with single allelic carriers as a single combined variant genotype (*AA/GG+AG/GA*) the frequency distribution was almost similar for both cases (78.2%) and controls (79.2%). No association towards lung cancer was observed when risk analysis (OR=0.95, 95%C.I.=0.49-1.87, p=0.91) was evaluated after logistic regression analysis.

Further, for the next combination of *XRCC1 Arg*³⁹⁹ *Gln* and *His*⁶³²*His*, lung cancer subjects carrying the mutant alleles (*AA/AA*) for both SNPs had a higher representation as compared to controls (21.7% vs 11.1%). Taking wild type (*GG/GG*) genotype for both SNPs as a reference a trend towards a high risk for lung cancer was observed (OR=2.13, 95%C.I.=0.79-5.7, p=0.13).

As shown in table 5.18 for the next double combination of *XRCC1 Arg*³⁹⁹ *Gln* and *Arg*²⁸⁰*His*, risk analysis similar to the above combination suggested a trend towards high risk for mutant genotype (*AA/AA*). Lung cancer subjects with mutant genotype for both the *XRCC1* SNPs (5.5%) had a higher representation compared to controls (2.5%). Nine-fold risk was observed towards lung cancer by double allelic variant subjects for the combination (OR=8.71, 95%C.I.=0.59-128.65, p=0.11).

Furthermore, as shown in table 5.18 it was observed subjects having single copy of mutant allele for both *XRCC1 Arg*¹⁹⁴*Trp* and *Pro*²⁰⁶*Pro* (*CT/GA*) were found to be more

represented in cases (23.5%) as compared to controls (16.9%). Taking the wild type genotype for both SNPs as a reference a trend of high risk was observed (*CC/GG*) (OR=1.43, 95%C.I.=0.76-2.67, $p=0.25$). On the same lines, for the combination of *XRCC1 Arg¹⁹⁴Trp* and *His⁶³²His* the percentage of heterozygous subjects (*CT/GA*) for both the SNPs was higher in cases (30.3%) when compared to the controls (17.2%). Taking wild type genotype (*CC/GG*) as a reference for both SNPs, a significant two-fold risk was observed in lung cancer subjects (OR=2.02, 95%C.I.=0.80-2.77, $p=0.04$, FDR_p=0.2) as shown in table 5.18.

In case of *XRCC1 Pro²⁰⁶Pro* and *XRCC1 Gln⁶³²Gln* there were 80.6% of subjects in the diseased group carrying single allelic (*AG/GA*) variant genotypes for both *XRCC1 Pro²⁰⁶Pro* and *Gln⁶³²Gln* SNPs. Lung cancer patients carrying the mutant alleles for both the polymorphic sites of *XRCC1* i.e. *XRCC1 Pro²⁰⁶Pro* and *XRCC1 Gln⁶³²Gln* had a six fold risk towards lung cancer as compared to subjects homozygous for the wild-type allele for both SNPs (*AA/GG*) (OR=6.69, 95%C.I.=3.56-12.58, $p<0.0001$, FDR_p=**0.001**). When dichotomous (*AG/GG+GA/GG*) group was analyzed for the above discussed combination a similar elevated risk towards lung cancer was observed (OR=5.6, 95%C.I.=3.0-10.4, $p<0.0004$, FDR_p=0.123) as shown in table 5.18.

Table 5.18: Genotypic distribution based on different double genotypic combinations of *XRCC1* variants and their association with lung cancer risk

Genotype	Cases, n (%)	Controls, n (%)	Overall AOR (95% CI) ^a	p ^b	FDR _p ^c
<i>XRCC1 Arg³⁹⁹Gln, G/A & XRCC1 Arg¹⁹⁴Trp, C/T</i>	N=111	N = 97			
GG/CC	68 (61.2)	64 (65.9)	1.00 (Reference)		
GA/CT	42 (37.8)	33 (34.0)	1.135 (0.621-2.0)	0.67	0.922
AA/TT	1 (0.9)	-	-	-	
GA+AA/CT+TT	43 (38.7)	33 (34.0)	1.17 (0.64-2.13)	0.60	0.911
<i>XRCC1 Arg³⁹⁹Gln, G/A & XRCC1 Pro²⁰⁶Pro, A/G</i>	N=124	N=106			
GG/AA	27 (21.7)	22 (20.7)	1 (Reference)		
GA/AG	93 (75)	82 (77.3)	0.93 (0.47-1.81)	0.83	0.922
AA/GG	4 (3.2)	2 (1.8)	2.17 (0.31-15.14)	0.43	0.537
GA+AA/AG+GG	97 (78.2)	84 (79.2)	0.96 (0.49-1.87)	0.91	0.91
<i>XRCC1 Arg³⁹⁹Gln, G/A & XRCC1 Gln⁶³²Gln, G/A</i>	N=124	N=108			
GG/GG	21 (16.9)	18 (16.6)	1.00 (Reference)		
GA/GA	76 (61.2)	78 (72.2)	0.80 (0.38-1.67)	0.55	0.916
AA/AA	27 (21.7)	12 (11.1)	2.13 (0.79-5.7)	0.13	0.325

to be cont'd on next page

GA+AA/GA+AA	103 (83.0)	90 (83.3)	0.95 (0.46-1.96)	0.90	0.911
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC1 Arg ²⁸⁰ His, G/A	N=162	N=155			
GG/GG	10 (6.1)	8 (5.1)	1.00 (Reference)		
GA/GA	143 (88.2)	143 (92.2)	0.86 (0.32-2.30)	0.18	0.6
AA/AA	9 (5.5)	4 (2.5)	8.71 (0.59-128.65)	0.11	0.325
GA+AA/GA+AA	152 (93.8)	147 (94.8)	0.90 (0.34-2.40)	0.84	0.91
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC1 Pro ²⁰⁶ Pro, A/G	N=123	N=142			
CC/AA	93 (75.6)	118 (83.0)	1.00 (Reference)		
CT/AG	29 (23.5)	24 (16.9)	1.43 (0.76-2.67)	0.25	0.625
TT/GG	1 (0.81)	-	-	-	
CT+TT/AG+GG	30 (24.3)	24 (16.9)	1.49 (0.80-2.77)	0.20	0.91
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC1 Gln ⁶³² Gln, G/A	N=79	N=116			
CC/GG	54 (68.3)	96 (82.7)	1.00 (Reference)		
CT/GA	24 (30.3)	20 (17.2)	2.02 (1.00-4.07)	0.04	0.2
TT/AA	1 (1.2)	-	-	-	
CT+TT/GA+AA	25 (31.6)	20 (17.2)	2.10 (1.05-4.20)	0.03	0.533
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC1 Arg ²⁸⁰ His, G/A	N=76	N=63			
CC/GG	22 (28.9)	19 (30.1)	1.00 (Reference)		
CT/GA	54 (71.0)	44 (69.8)	1.1 (0.52-2.30)	0.77	0.922
XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC1 Gln ⁶³² Gln, G/A	N=124	N=188			
AA/GG	16 (12.9)	87 (46.2)	1.00 (Reference)		
AG/GA	100 (80.6)	76 (40.4)	6.69 (3.56-12.58)	<0.0001	0.001
GG/AA	8 (6.4)	25(13.2)	1.76 (0.67-4.60)	0.25	0.417
AG+GG/GA+AA	108 (87.0)	101 (53.7)	5.6 (3.00-10.40)	<0.0004	0.120
XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC1 Arg ²⁸⁰ His, G/A	N=135	N=134			
AA/GG	11 (8.14)	15 (11.1)	1.00 (Reference)		
AG/GA	119 (88.1)	114 (85.0)	1.50 (0.64-3.49)	0.34	0.68
GG/AA	5 (3.7)	5(3.7)	1.47 (0.29-7.29)	0.63	0.63
AG+GG/GA+AA	124 (91.8)	119 (88.8)	1.5 (0.65-3.50)	0.32	0.0033

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. (c)- False discovery rate. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 3: combined hetero and mutant genotype 4:FDR - False Discovery Rate Adjusted p

5.6.7 Combinatorial analysis of risk associated with triple combinations of different polymorphic sites of *XRCC1* gene

Furthermore, in the current study we also evaluated the triple combinations for *XRCC1* polymorphic sites for having any inflection towards lung cancer susceptibility. Subjects who carried a single variant allele for the three SNPs *i.e.* (*XRCC1* Arg³⁹⁹Gln/Pro²⁰⁶Pro/ Gln⁶³²Gln (*GA/AG/GA*)) were found to be at higher incidence in lung cancer patients as compared to control subjects (90.1% vs 68.8%) ($\chi^2=3.8$, $p=0.14$, $df=2$). This combination depicted a risk of 2.7 fold towards lung cancer development (OR=2.7, 95%C.I.=0.86-8.96, $p=0.08$) as compared to subjects with wild type (*GG+AA+GG*) genotype. When both the heterozygous and mutant genotypes were combined as a single genotype (*GA+AA/AG+GG/GA+AA*) it was observed that 91.8% cases in the above combination were found to be more represented in lung cancer cases as compared to 80.3% of controls. An almost three-folds risk was observed, though statistically insignificant (OR=2.7, 95%C.I.=0.84-8.80, $p=0.09$) as table 5.19 shows.

The triple combination comprising of (*Arg*¹⁹⁴*Trp*, *Pro*²⁰⁶*Pro* and *Gln*⁶³²*Gln*) of *XRCC1* genotypes revealed that 55.1% of the lung cancer subjects carried a single copy of the variant allele (*CT/AG/GA*) for the three polymorphic sites of *XRCC1* gene as compared to 16.0% for controls ($\chi^2=14.8$, $p=0.0001$, $df=2$). Whereas, there were 83.9% subjects with wild type genotype (*CC/AA/GG*) in controls as compared to 44.8% in the diseased group. Logistic regression analysis showed that individuals with single variant genotypic combination for the above three alleles were at 6-fold risk for lung cancer as compared to those who had no variant allele for all the three SNPs and was found to be significant (OR=6.1, 95%C.I.=2.34-16.2, **$p=0.0002$**).

Table 5.19: Genotypic distribution based on triple genotypic combinations of different *XRCC1* variants and their association with lung cancer risk

Genotype	Cases, n (%)	Controls, n (%)	Overall AOR (95% CI) ^a	p ^b	FDR _p ^c
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC1 Pro ²⁰⁶ Pro, A/G	N=29	N=30			
GG/CC/AA	17 (79.9)	18 (88.3)	1.00 (Reference)		
GA/CT/AG	11 (20.1)	12 (11.7)	0.96 (0.32-2.90)	0.95	0.96
AA/TT/GG	1 (3.4)	-	-	-	-
GA+AA/CT+TT/AG+GG	12	12	1.03 (0.36-3.18)	0.89	0.89
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC1 Gln ⁶³² Gln, G/A	N=61	N=61			
GG/AA/GG	5 (8.1)	12 (19.6)	1.00 (Reference)		
GA+AG+GA	55 (90.1)	47 (77.0)	2.7 (0.86-8.96)	0.087	0.21
AA/GG/AA	1 (1.6)	2 (3.2)	-	-	-
GA+AA/AG+GG/GA+AA	56 (91.8)	49 (80.3)	2.7 (0.84-8.80)	0.09	0.18
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC1 Arg ²⁸⁰ His, G/A	N=72	N=69			
GG/AA/GG	4 (5.5)	3 (4.3)	1.00 (Reference)		
GA/AG/GA	68 (94.4)	66 (95.6)	0.96 (0.20-4.60)	0.96	0.96
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC1 Arg ²⁸⁰ His, G/A	N=37	N=33			
GG/CC/GG	6 (16.2)	6 (18.8)	1.00 (Reference)		
GA/CT/GA	31 (83.7)	27 (81.8)	1.37 (0.38-4.90)	0.62	0.96
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC1 Gln ⁶³² Gln, G/A	N=29	N=81			
CC/AA/GG	13 (44.8)	68 (83.9)	1.00 (Reference)		
CT/AG/GA	16 (55.1)	13 (16.0)	6.1 (2.34-16.20)	0.0002	0.001

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. (c)-False discovery rate p-value. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 3: combined hetero and mutant genotype

5.6.8 Haplotypes and Linkage disequilibrium analysis of *XRCC1* variants and their association with lung cancer susceptibility

The five polymorphic variants of the *XRCC1* gene are present on the same gene therefore to look at genetic association of varied SNPs, linkage disequilibrium, haplotype analysis was done. SHEsis, program was used to carry out the haplotype analysis. Haplotype

Table 5.20: Analyzed Haplotype frequencies and risk towards lung cancer due to *XRCC1* haplotypes

HAPLOTYPE	HAPLOTYPE CONSTRUC- TION (194-206- 280-399-632)	Case fre- quency, n (%)	Control fre- quency, n (%)	AOR (95% CI) ^a	p ^b
Hap1	CAAAA	0.125	0.052	2.69 (1.76-4.08)	2.06e-006
Hap2	CAGAA	0.122	0.038	3.63 (2.27-5.79)	1.18e-008
Hap3	CAAAG	0.063	0.140	0.42 (0.28-0.61)	5.26e-006
Hap4	CAGAG	0.007	0.143	0.04 (0.02-0.11)	1.14e-020
Hap5	CAAGA	0.064	0.031	2.17 (1.26-3.74)	0.0042
Hap6	CAGGA	0.089	0.042	2.28 (1.43-3.65)	0.0004
Hap7	CAAGG	0.041	0.060	0.68 (0.41-1.12)	0.1259
Hap8	CAGGG	0.040	0.087	0.45 (0.28-0.72)	0.0006
Hap9	CGAGA	0.031	0.136	0.20 (0.13-0.34)	8.47e-012
Hap10	CGGGA	0.043	0.058	0.73 (0.44-1.21)	0.2188
Hap11	CGAGG	0.103	0.019	6.12 (3.29-11.37)	1.48e-010
Hap12	CGGGG	0.068	0.024	3.05 (1.69-5.50)	0.0001
Hap13	TAGGG	0.011	0.039	0.29 (0.13-0.65)	0.0015

(a)-Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by logistic regression analysis. (b)-Two-sided χ^2 test for either haplotypic distribution or frequencies between the cases and controls

frequencies and linkage disequilibrium were calculated for the five SNPs. Haplotype frequencies are tabulated in table 5.20. Out of all 30 haplotypes, those with cases and control frequency <0.03 were omitted from the analysis. Comparison of overall haplotype profile revealed a significant difference between the cases and controls (global test: $p=0.00e+00$). As shown in table 5.20 the haplotype block 11 (*CGAGG*) was associated with the highest risk towards lung cancer (OR=6.12, 95%CI=3.28-11.37, $p=1.48e-010$). In the present study both the haplotype blocks 6 and 12 each carrying at least one mutant allele were shown to have a high risk for lung cancer *i.e.* haplotype block 6 (OR=2.281, $p=0.0004$) and haplotype block 12 (OR=3.054, $p=0.0010$). The haplotype blocks with wild allele for all positions (*CAGGG*) depicted a strong protective effect

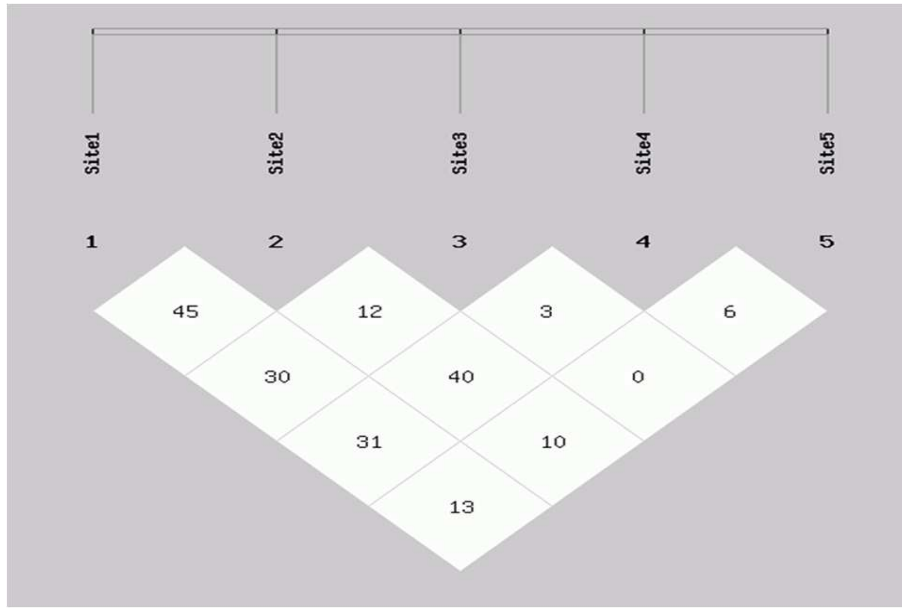


Figure 5.14: Pairwise linkage disequilibrium (D') between the five variants of *XRCC1* gene. White color depicts the minimum LD

Table 5.21: D'/r^2 Values for linkage disequilibrium among five *XRCC1* SNPs

CODON	194	206	280	399	632
194	-	0.45/0.012	0.30/0.012	0.32/0.010	0.136/0.002
206	-	-	0.13/0.008	0.401/0.071	0.10/0.005
280	-	-	-	0.36/0.001	0.002/0.00
399	-	-	-	-	0.063/0.004
632	-	-	-	-	-

towards lung cancer (OR=0.44, 95%C.I.=0.277-0.717, $p=0.000674$). Haplotype block 13 (*TAGGG*) with mutant allele for *XRCC1 Arg*¹⁹⁴*Trp* and wild type allele for the rest of the *XRCC1* SNPs showed a strong protective effect towards lung cancer (OR=0.29, 95%C.I.=0.13-0.65, $p=0.0015$). Similarly haplotype block-4 (*CAGAG*) carrying minor allele for *XRCC1 Arg*³⁹⁹*Gln* and wild type for rest of the *XRCC1* SNPs had the strongest protective effect (OR=0.044, 95%C.I.=0.02-0.11, $p=1.14e-020$) towards lung cancer.

Table 5.21 summarizes the result of D values and r^2 values for linkage dis-equilibrium (LD) between markers for cases and controls together. *XRCC1 Arg*¹⁹⁴*Trp* gave a linkage disequilibrium with *XRCC1 Pro*²⁰⁶*Pro* ($D=0.45$, $r^2=0.012$). Similarly, *XRCC1 Arg*³⁹⁹*Gln* and *XRCC1 Arg*¹⁹⁴*Trp* illustrated a linkage disequilibrium ($D=0.320$, $r^2=0.010$). However, *XRCC1 Gln*¹⁹⁴*Gln* gave a weak LD with *XRCC1 Arg*²⁸⁰*His* ($D=0.002$, $r^2=0.00$) as shown in figure 5.14.

5.6.9 Multifactor Dimensionality Reduction (MDR) analysis for *XRCC1* variants

Traditional logistic regression approach, lacks in the evaluation of high-order gene-gene interactions. Therefore, MDR is used to determine high order gene-gene interactions. MDR is a non-parametric method used to determine the major factors governing the lung cancer susceptibility in North Indian population. Table 5.22 summarizes the average cross validation consistency (CVC) and average prediction error obtained from MDR analysis of the data set of subjects with or without lung cancer. Our analysis demonstrated that the best interaction model was the three factor model which included SNPs *XRCC1* 206, 632, 280 having a CVC of 10 on 10 and prediction error of 0.34 (maximum CVC, minimum prediction error). *XRCC1* *Gln*⁶³²*Gln* acted as the best one-factor model for predicting

Table 5.22: Multifactor Dimensionality reduction (MDR) analysis of *XRCC1* variants showing association of high order interactions with Lung cancer

No. of risk factors	Best interaction Model	Cross Validation Consistency (CVC)	Prediction Error	p-value
1	<i>XRCC1</i> 632	10/10	0.41	0.07
2	<i>XRCC1</i> 206, <i>XRCC1</i> 632	10/10	0.39	0.06
3	<i>XRCC1</i> 206, <i>XRCC1</i> 632, <i>XRCC1</i> 280	10/10	0.34	0.01
4	<i>XRCC1</i> 399, <i>XRCC1</i> 206, <i>XRCC1</i> 632, <i>XRCC1</i> 280	8/10	0.38	0.04
5	<i>XRCC1</i> 399, <i>XRCC1</i> 1194, <i>XRCC1</i> 206, <i>XRCC1</i> 632, <i>XRCC1</i> 280	10/10	0.36	0.02

lung cancer risk with CVC of 10 and prediction error of 0.41. The two factor model included *XRCC1* *Pro*²⁰⁶*Pro* and *XRCC1* *Gln*⁶³²*Gln*, it had an improvised prediction error of 0.39 with CVC of 10. The CVC for the four factor model factor was less as compared to other model but an improvised prediction error of 0.38. The best five interaction model had all 5 variants with CVC of 10 and prediction error of 0.36 as shown in table 5.22.

The interaction dendrogram for the cluster analysis is given in figure 5.15. Colored lines describe the type of interaction between the interacting SNPs. Red and orange indicate synergy between SNPs. Yellow suggests independence. Green and blue indicate redundancy or no interaction. The shorter the lines connecting the two SNPs stronger the interaction. Synergy exist between *XRCC1* *Pro*²⁰⁶*Pro* and *XRCC1* *Gln*⁶³²*Gln* as they

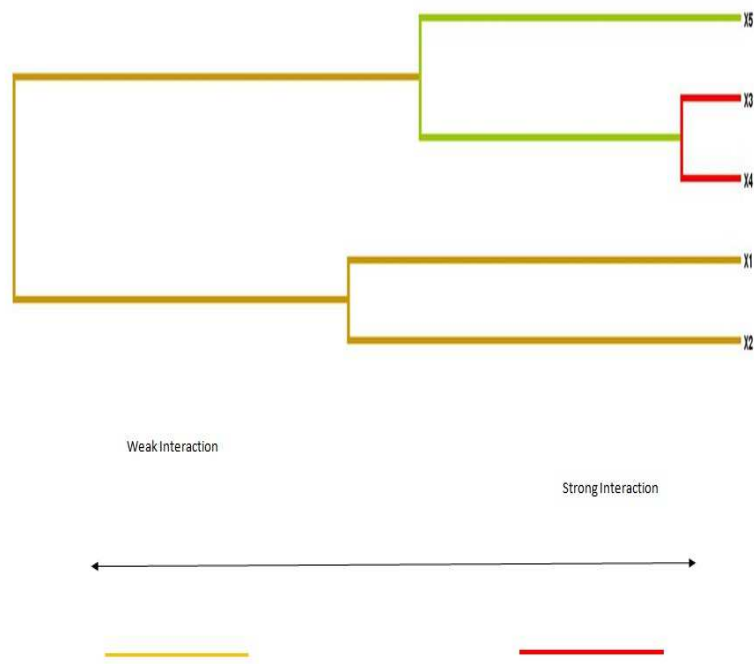


Figure 5.15: The multifactor dimensionality reduction (MDR) interaction dendrogram. *X1*, *XRCC1* 399; *X2*, *XRCC1* 194; *X3*, *XRCC1* 206; *X4*, *XRCC1* 632; *X5*, *XRCC1* 280

had the smallest line connecting them. Further *XRCC1* Arg²⁸⁰His/ *XRCC1* Pro²⁰⁶Pro and *XRCC1* Gln⁶³²Gln are on same line, indicating strong interaction exists between them, whereas *XRCC1* Arg³⁹⁹Gln and *XRCC1* Arg¹⁹⁴Trp are independent of each other. However, analyzing all SNPs together there is a weak interaction between the five as shown in figure 5.15

5.6.10 Classification and Regression Tree (CART) analysis for five polymorphic variants of *XRCC1* gene

High order SNP-SNP interactions were explored using CART analysis. The final tree generated is shown in figure 5.16. The tree consists of 9 terminal nodes as shown in table 5.23. The root split was on *XRCC1* Gln⁶³²Gln which was consistent with MDR best one factor model, suggesting that this SNP is the strongest risk factor for lung cancer among the interacting variants. Subsequent split were on *XRCC1* Pro²⁰⁶Pro and *XRCC1* Arg²⁸⁰His, *XRCC1* Arg¹⁹⁴Trp, *XRCC1* Arg²⁰⁶Gln respectively as shown in different nodes in figure 5.16.

To calculate ORs we chose node 9 as reference with minimum case rate of 15.5% and lowest risk as shown in table 5.23. Node 9 represented individuals carrying the wild genotype of *XRCC1 Pro*²⁰⁶*Pro* and *XRCC1 Gln*⁶³²*Gln*. Individuals carrying both the *XRCC1 Arg*²⁸⁰*His* (W) and *XRCC1 Gln*⁶³²*Gln* (M) exhibited the highest risk towards lung cancer which was 29 folds (OR=29.0, 95%C.I.=10.43-80.5, ***p*=0.0001**). Node-5 carrying individuals with combined genotype of *XRCC1 Arg*³⁹⁹*Gln* (W)/ *XRCC1 Arg*²⁸⁰*His* (M)/*XRCC1 Pro*²⁰⁶*Pro* (M)/*XRCC1 Gln*⁶³²*Gln* (W) also expressed a very high risk towards lung cancer (OR=21.75, 95%C.I.=6.43-73.55, ***p*=0.0001**). Interacting group involving the five polymorphic variants with genotypic combination of *XRCC1 Arg*¹⁹⁴*Trp* (W)/ *XRCC1 Arg*³⁹⁹*Gln* (M)/ *XRCC1 Arg*²⁸⁰*His* (M)/ *XRCC1 Pro*²⁰⁶*Pro* (M)/ *XRCC1 Gln*⁶³²*Gln* (W) conferred a significant 8 fold-risk towards lung cancer (node6) (OR=8.15, 95%C.I.=3.66-18.55, ***p*=0.0001**). However, replacing *XRCC1 Arg*¹⁹⁴*Trp* wild types with *XRCC1 Arg*¹⁹⁴*Trp* mutant (heterozygotes+mutants) genotype lowered the risk to 1.55 (OR=1.55, 95%C.I.=0.29-8.16, *p*=0.602) (node7). Subjects with genotypic combination which excluded *XRCC1 Arg*³⁹⁹*Gln* i.e. *XRCC1 Gln*⁶³²*Gln* (M)/ *Pro*²⁰⁶*Pro* (M)/ *Arg*²⁸⁰*His* (M) / *Gln*⁶³²*Gln* (M) or *XRCC1 Gln*⁶³²*Gln*(W)/ *Pro*²⁰⁶*Pro* (M)/ *Arg*²⁸⁰*His* (M)/ *Gln*⁶³²*Gln* (M) showed a significant risk towards lung cancer as shown in table 5.23.

Table 5.23: Risk estimates based on CART analysis for *XRCC1* variants

Node	Genotype in each terminal node	Controls (n)	Cases (n)	Case Rate ^a	OR (95% CI)	p ^b
Node 1	XRCC1 280(W)/XRCC1 632(M)	6	32	84.2	29.0 (10.43-80.50)	0.0001
Node 2	XRCC1 206(W)/XRCC1 280(M)/XRCC1 632(M)	57	107	65.2	10.20 (5.47-19.02)	0.0001
Node 3	XRCC1 194(M)/XRCC1 206(M)/XRCC1 280(M)/XRCC1 632(M)	19	22	53.6	6.29 (2.79-14.19)	0.00001
Node 4	XRCC1 194(W)/XRCC1 206(M)/XRCC1 280(M)/XRCC1 632(M)	119	108	47.5	4.93 (2.72-8.90)	0.0001
Node 5	XRCC1 399(W)/XRCC1 280(M)/XRCC1 206(M)/XRCC1 632(W)	4	16	80	21.75 (6.43-73.55)	0.0001
Node 6	XRCC1 194(W)/XRCC1 399(M)/XRCC1 280(M)/XRCC1 206(M)/XRCC1 632(W)	18	27	60	8.15 (3.66-18.15)	0.0001
Node 7	XRCC1 194(M)/XRCC1 399(M)/XRCC1 280 (M)/XRCC1 206(M)/XRCC1 632(W)	7	2	22.2	1.55 (0.29-8.16)	0.602
Node 8	XRCC1 280(W)/XRCC1 206(M)/XRCC1 632(W)	8	0	-	0.00 (-)	-
Node 9	XRCC1 206(W)/XRCC1 632(W)	87	16	15.5	1.0 ^c	-

(a)-case rate is percentage LC cases among all individuals in each terminal node case/(case+ control) x100; (b)-p<0.05 is considered statistically significant; c reference group; no adjustment for age, gender and pack years is done

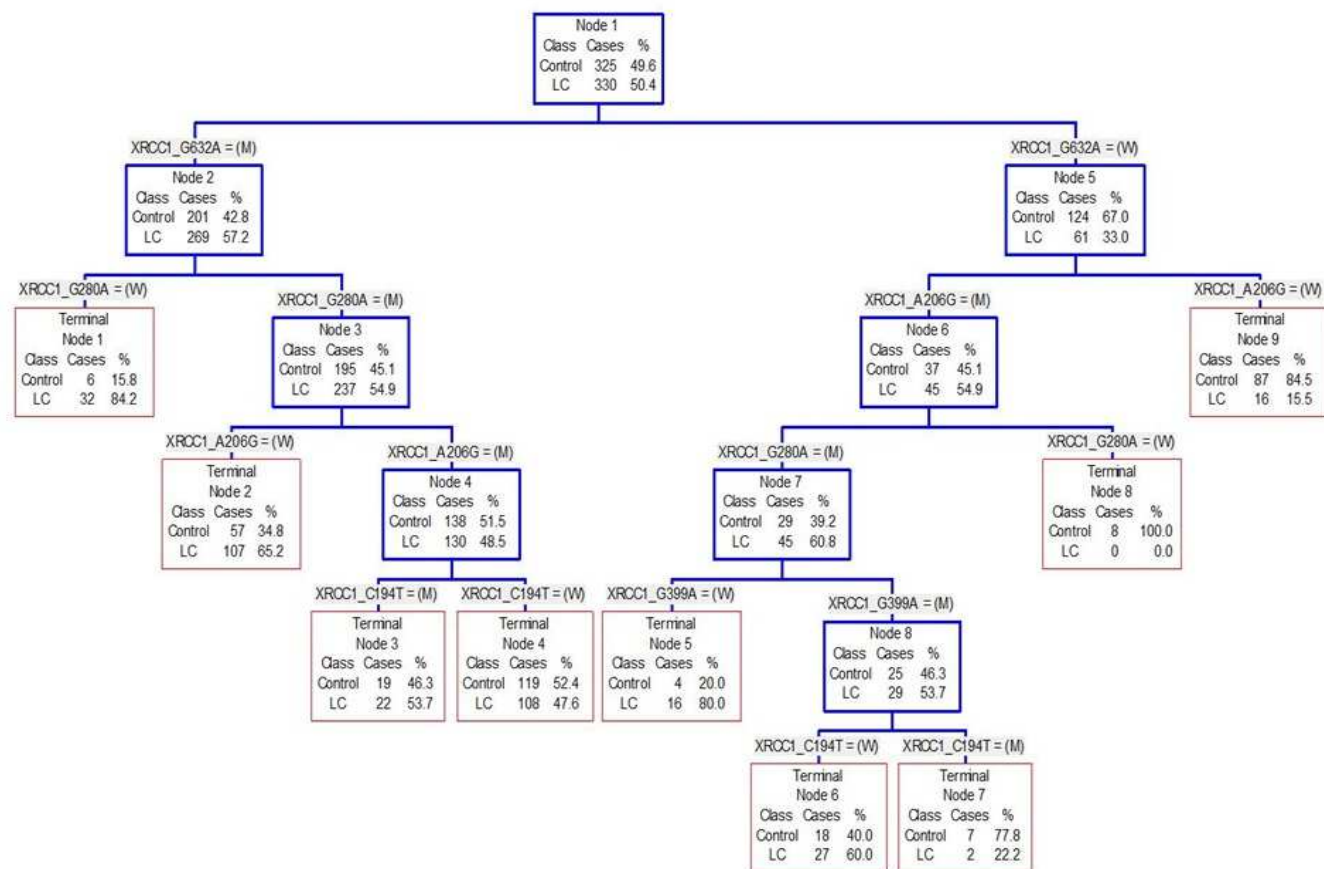


Figure 5.16: CART analysis in polymorphic *XRCC1*. W= homozygous wild type genotype, M= heterozygous + homozygous variant genotype

5.7 Relationship of polymorphic variants of DSB repair genes *XRCC6*, *XRCC7* in lung cancer susceptibility

5.7.1 Demography of lung cancer patients and controls with polymorphic *XRCC6* and *XRCC7*

The demographic details of patients and control subjects who were studied for both *XRCC6* and *XRCC7* SNP is detailed in table 5.24. The study enrolled 330 lung cancer cases and 320 healthy controls. The mean age for the diseased group was 58.11 ± 10.85 and 52.2 ± 10.46 ($p < 0.0001$) for the control group. The frequency distribution for gender is comparable as shown in table 5.24 and no significant difference was observed ($p = 0.348$). Diseased group had 86.6% (286) and 13.3% (44) males and females respectively while 83.7% (268) males and 16.2% (52) females were present in controls. When smoking was observed as a parameter there was a high representation of ever smokers (81.2%) in cases as compared to the controls (70%); whereas frequency of non-smokers was less as compared to controls (18.7% and 30%) and these differences were statistically significant ($p = 0.0001$). Similarly the mean value of pack years was higher for cases when compared to controls (28.04 ± 33.9 vs 17.16 , $p < 0.0001$).

Histological distribution revealed, there were 41.8% (138) squamous cell carcinoma (SQCC) patients, 32.4% (107) adenocarcinomas (ADCC), 24.2% (80) small cell carcinoma (SCLC) patients and remaining 1.5% (5) were unclassified. TNM staging assessment classified 0.9% (3) patients as stage I, 4.2% (14) as stage II, 48.7% (161) as stage III, 43.0% (142) as stage IV patients and 2.7% (9) were unclassified. The TNM classification was further detailed where subjects were classified for primary tumor extension, lymph node extension and metastasis. The primary tumor extension is classified from T0 to T4, the presence and extent of lymph node involvement is classified as N0, N1 till N3 and distant metastasis is classified into M0 and M1. The distribution of patients according to tumor size, lymph node, invasion and distant metastasis has been recorded in table 5.24. Distribution according to primary tumor extension comprised of Tx: 3.3% (11), T1: 5.1% (17), T2: 12.1% (40), T3: 26.3% (87), T4: 46.6% (154). According to lymph node involvement parameter, the patients are distributed as N0: 15.7% (52), N1: 11.2% (37), N2: 39.3% (130), N3: 27.2% (90). Also, the proportion of patients who did not show metastasis was 53.9% (178), as compared to 39.6% (131) who were showing positive metastasis. In all these three categories 10% of patients were classified as unknown.

Table 5.24: Demographic characteristics among cases and controls

Variable	Cases N=330 (n%)	Controls N=320 (n%)	p ^a
Age (Years)			
Mean±SD Range	52.22±10.46	58.11±10.85	< 0.0001
Gender			
Male	286 (86.6)	268 (83.7)	0.348
Female	44 (13.3)	52 (16.2)	
Smoking status			
Smokers	268 (81.2)	224 (70.0)	0.0012
Non-smokers	62 (18.7)	96 (30.0)	
Pack Years			
Mean±SD	28.04±33.9	17.16±19.8	< 0.0001
Histology			
ADCC	107 (32.4)		
SQCC	138 (41.8)		
SCLC	80 (24.2)		
Others	5 (1.5)		
TNM staging			
I	3 (0.9)		
II	14 (4.2)		
III	161 (48.7)		
IV	142 (43.0)		
Unclassified	10 (3.0)		
Tumor			
Tx	11 (3.3)		
T1	17 (5.1)		
T2	40 (12.1)		
T3	87 (26.3)		
T4	154 (46.6)		
Unknown	21 (6.3)		
Lymph node			
N0	52 (15.7)		
N1	37 (11.2)		
N2	130 (39.3)		
N3	90 (27.2)		
Unknown	21 (6.3)		
Metastasis			
M0	178 (53.9)		
M1	131 (39.6)		
Unknown	21 (6.3)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects, (a)-p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

5.7.2 Genotypic and allelic distribution of polymorphic *XRCC6* and *XRCC7* variants and their association with lung cancer risk

Table 5.25 shows the allelic and genotypic distribution of *XRCC6* 61C>G and *XRCC7* 6721G>T polymorphisms and their association towards lung cancer. The genotypic frequencies of these two polymorphisms were in agreement with the Hardy Weinberg equilibrium ($\chi^2=2.82$; $p=0.10$ for *XRCC6*-61C>G) and ($\chi^2=0.36$, $p=0.54$ for *XRCC7* 6721G>T). For *XRCC6* 61C>G polymorphism, the frequencies of the wild (CC), heterozygous (CG), and mutant (GG) genotypes were 43.6%, 47.8%, and 8.4% among the cases and 46.2%, 50.3%, and 3.4%, respectively, among the controls. Therefore a significant difference in the genotypic distribution between cases and controls ($\chi^2=6.63$, $df=2$, $p=0.03$) for *XRCC6* 61C>G was observed. The minor allele frequency for the *XRCC6* 61C>G was higher in cases as compared to controls (0.32 vs.0.29). As shown in table 5.25, logistic regression analysis revealed that there was a lack of significant association of *XRCC6* 61G>C polymorphism towards risk for lung cancer. However carriers for double allelic variant (GG) genotype, depicted a two fold risk towards lung cancer (OR=1.97, 95%C.I.=0.90-1.43, $p=0.08$) with approaching significance. Subjects harbouring single copy of variant genotype (CG) (OR=0.89, 95%C.I.=0.63-1.24, $p=0.50$) and combined variant (CG+GG)(OR=0.96, 95%C.I.=0.69-1.33, $p=0.83$) genotypes showed no association towards lung cancer.

For, *XRCC7* 6721G>T polymorphism, 18.4% of lung cancer subjects carried wild type genotype (GG), 50.9% were single allelic variant carriers (GT) and 30.6% subjects harbored mutant (TT) genotype. In case of controls there were 21.2% wild type subjects, 49.6% heterozygous, and 29% mutant subjects, respectively. No difference in the genotypic distribution between cases and controls ($\chi^2=0.80$, $df=2$, $p=0.66$) for *XRCC7* 6721G>T polymorphism was observed. Frequency of single allelic carriers (GT) was similar in cases and controls (50.9% vs 49.6%). No association was observed towards lung cancer as shown in table 5.25. Similarly mutant genotype (TT) showed no risk (OR=1.23, 95%C.I.=0.77-1.96, $p=0.36$) towards lung cancer as the frequency distribution was almost similar (30.6% vs 29.0%) as shown in table 5.25

Table 5.25: Genotypic distribution of *XRCC6* and *XRCC7* genetic variants and their association with risk of lung cancer

XRCC6 61C>G	Controls N=320 (n%)	Cases N=330 (n%)	Overall AOR (95% CI)^a	p^b
CC	148 (46.2)	144 (43.6)	1	
CG	161 (50.3)	158 (47.8)	0.89 (0.63-1.24)	0.50
GG	11 (3.4)	28 (8.4)	1.97 (0.90-4.31)	0.08
CG+GG	172 (53.7)	186 (56.3)	0.96 (0.69-1.33)	0.83
C	457	446		
G	183	214		
MAF	0.29	0.32		
XRCC7 6721G>T	Controls N=320 (n%)	Cases N=330 (n%)	Overall AOR (95% CI)^a	p^b
GG	68 (21.2)	61 (18.4)	1	
GT	159 (49.6)	168 (50.9)	1.08 (0.70-1.65)	0.72
TT	93 (29.0)	101 (30.6)	1.23 (0.77-1.96)	0.36
GT+TT	252 (78.7)	269 (81.5)	1.14 (0.76-1.70)	0.52
G	295	290		
T	345	370		
MAF	0.53	0.56		

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

5.7.3 Genotypic and allelic distribution of *XRCC6* and *XRCC7* genetic variants and the stratified association analysis based on histological subtypes

In order to stratify the study subjects on the basis of histology into high risk subgroups the lung cancer subjects were evaluated for the three histologies *i.e.* ADCC, SQCC, and SCLC as shown in table 5.26. In case of *XRCC6 61C>G* subjects with ADCC subtype and carrying mutant genotype (*GG*) had a higher representation in cases (8.44%) as compared to controls (3.4%), however, the difference was not significant and a visible trend of risk was observed towards lung cancer (OR=1.48, 95%C.I.=0.51-4.29, p=0.46). Further, the frequency of single allelic carriers (*CG*) was similar in both SQCC (49.2%) and SCLC (51.2%) subgroups when compared to controls (50.3%). On taking wild type (*CC*) genotype as a reference no affect on lung cancer was observed for both SQCC (OR=0.98, 95%C.I.=0.63-1.53, p=0.95) and SCLC (OR=0.96, 95%C.I.=0.55-1.67, p=0.90) subtypes,

as shown in table 5.26. Furthermore, study subjects with double allelic variant or mutant (*GG*) genotype in case of both SQCC (8.6%) and SCLC (8.75%) patients showed a higher frequency as compared to controls (3.4%). On comparing with the wild type (*CC*) genotype a visible trend of 2-fold risk for both SQCC (OR=2.03, 95%C.I.=0.77-5.33, $p=0.14$) and SCLC (OR=2.2, 95%C.I.=0.73-6.69, $p=0.15$) subtype was observed as shown in table 5.26.

For *XRCC7 6721G>T* polymorphism, the stratified analysis revealed, no association for ADCC and SQCC subtype, whereas SCLC subjects showed an elevated risk for all the genotypic combinations. The frequency distribution in case of lung cancer subjects diagnosed with ADCC subtype, depicted there were 50.4% subjects with heterozygous genotype (*GT*) and it was almost similar to that of controls (49.6 %). As shown in table 5.26 logistic regression analysis revealed an OR of 1 which depicted no association with lung cancer. In case of subjects carrying mutant genotype (*TT*) there was a higher representation of frequency of ADCC subjects (29.0%) compared to controls (25.2%) however, no association was observed towards lung cancer risk (OR=0.82, 95%C.I.=0.43-1.58, $p=0.57$) as stated in table 5.26.

Lung cancer diseased group subjects with SQCC subtype and carrying single allele of variant genotype (*GT*) had similar representation to controls (50.7% vs 49.6%). Taking wild type (*GG*) subjects as a reference no risk was observed towards lung cancer (OR=0.90, 95%C.I.=0.51-1.59, $p=0.74$). SCLC Subjects with mutant genotype (*TT*) had a higher frequency (41.2%) when compared to controls (29.0%). When compared with wild type (*GG*) genotype, the risk analysis showed a significant six fold risk for SCLC subtype (OR=6.41, 95%C.I.=2.08-19.7, **$p=0.001$**). On the same lines SCLC subjects carrying heterozygous genotype (*GT*) (OR=3.54, 95%C.I.=1.27-9.85, **$p=0.015$**) showed an elevated risk towards lung cancer. Combining both heterozygous and mutant genotype (*GT+TT*) as a single group, four fold risk was observed (OR=4.11, 95%C.I.=1.57-10.78, **$p=0.004$**) towards lung cancer development.

Table 5.26: Genotypic distribution of *XRCC6* and *XRCC7* genetic variants and their association with risk of lung cancer along with the stratified association based on histology

Genes		ADCC			SQCC			SCLC		
XRCC6 61C>G	Controls N=320 (n%)	Cases N=107 (n%)	AOR (95% CI)^a	p^b	Cases N=138 (n%)	AOR (95% CI)^a	p^b	Cases N=80 (n%)	AOR (95% CI)^a	p^b
CC	148 (46.2)	52 (48.5)	1.00 (Reference)		58 (42.0)	1.00 (Reference)		32 (40.0)	1.00 (Reference)	
CG	161 (50.3)	46 (42.9)	0.70 (0.63-1.24)	0.15	68 (49.2)	0.98 (0.63-1.53)	0.95	41 (51.2)	0.96 (0.55-1.67)	0.90
GG	11 (3.4)	9 (8.4)	1.48 (0.51-4.29)	0.46	12 (8.6)	2.03 (0.77-5.33)	0.14	7 (8.75)	2.2 (0.73-6.69)	0.15
CG+GG	172 (53.7)	55 (51.4)	0.77 (0.48-1.22)	0.27	80 (57.9)	1.06 (0.06-1.63)	0.77	48 (60.0)	1.06 (0.62-1.81)	0.82
XRCC7 6721G>T	Controls N=320 (n%)	Cases N=107 (n%)	AOR (95% CI)^a	p^b	Cases N=138 (n%)	AOR (95% CI)^a	p^b	Cases N=80 (n%)	AOR (95% CI)^a	p^b
GG	68 (21.2)	26 (24.2)	1.00 (Reference)		29 (21.0)	1.00 (Reference)		6 (7.5)	1.00 (Reference)	
GT	159 (49.6)	54 (50.4)	0.93 (0.52-1.64)	0.81	70 (50.7)	0.90 (0.51-1.59)	0.74	41 (51.2)	3.54 (1.27-9.85)	0.01
TT	93 (29.0)	27 (25.2)	0.82 (0.43-1.58)	0.57	39 (28.2)	1.04 (0.57-1.90)	0.89	33 (41.2)	6.41 (2.08-19.7)	0.001
GT+TT	252 (78.7)	81 (75.7)	0.90 (0.52-1.54)	0.70	109 (78.9)	0.96 (0.57-1.63)	0.90	74 (92.5)	4.11 (1.57-10.8)	0.004

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.7.4 Genotypic distribution of *XRCC6* and *XRCC7* genetic variants based on smoking status and its association towards risk of developing lung cancer

Gene-environment interaction was studied for both *XRCC6* and *XRCC7* SNPs' by classifying the study subjects into smokers and non-smokers as shown in table 5.27. In case of *XRCC6* 61C>G, smokers and lung cancer patients who were smokers and carrying heterozygous genotype (*CG*) had almost similar frequency distribution (51.1% vs 49.5%) in relation with controls who were smokers. On taking wild type genotype (*CC*) as a reference no association was observed for heterozygous subjects towards any association for lung cancer. Within non-smokers group 33.8% lung cancer subjects had heterozygous genotype (*CG*) who were non-smokers as compared to 52.0% of controls subjects. Logistic regression analysis exhibited a significant protective effect for heterozygous subjects (*CG*) (OR=0.38, 95%C.I.=0.18-0.82, ***p*=0.01**) towards lung cancer. Smokers carrying mutant genotype (*GG*) were found to be highly represented in cases as compared to controls (8.5% vs 3.1%). On comparison with the reference genotype (*CC*) a 1.6 fold risk (OR=1.6, 95%CI=1.02-2.55, ***p*=0.04**) was observed towards lung cancer for subjects carrying the mutant genotype (*GG*) and who were smokers. For non-smokers, frequency distribution exhibited there were 4.1% subjects with mutant genotype (*GG*) in controls and 8.0% subjects in cases. Association study depicted no risk towards lung cancer (OR=0.74, 95%C.I.=0.14-3.0, *p*=0.72) as shown in table 5.27.

In case of polymorphic *XRCC7* 6721G>T, for subjects who were smokers and carrying single allele of variant genotype (*GT*) had a higher representation in cases (53.3%) compared to controls (51.7%). However, the difference between the two was not significant therefore when compared to wild type (*GG*) genotype, no association towards lung cancer was observed (OR=1.22, 95%C.I.=0.73-1.99, *p*=0.44) as shown in table 5.27. Frequency of non-smokers with heterozygous genotype was 44.7% compared to 40.3% in cases. An odds ratio of 0.84 fold was obtained which signified no association (OR=0.84, 95%C.I.=0.35-1.99, *p*=0.69).

Table 5.27: Genotypic distribution of *XRCC6* and *XRCC7* based on smoking status and its association with risk of Lung cancer

Genotype	SMOKERS				NON SMOKERS			
	Controls N=224 (n%)	Cases N=268 (n%)	AOR (95% CI) ^a	p ^b	Controls N=96 (n%)	Cases N=62 (n%)	AOR (95% CI) ^a	p ^b
CC	106 (47.3)	108 (40.2)	1.00 (Reference)		42 (43.7)	36 (58)	1.00 (Reference)	
CG	111 (49.5)	137 (51.1)	1.11 (0.76-1.62)	0.58	50 (52.0)	21 (33.8)	0.38 (0.18-0.82)	0.01
GG	7 (3.1)	23 (8.5)	1.6 (1.02-2.55)	0.04	4 (4.1)	5 (8.0)	0.74 (0.14-3.0)	0.72
CG+GG	118 (52.6)	160 (59.7)	1.20 (0.83-1.76)	0.31	54 (56.2)	26 (41.9)	0.41 (0.20-0.86)	0.018
XRCC7 67 21G>T	Controls N=224 (n%)	Cases N=268 (n%)	AOR (95% CI)^a	p^b	Controls N=96 (n%)	Cases N=62 (n%)	AOR (95% CI)^a	p^b
GG	45 (20.0)	46 (17.1)	1.00 (Reference)		23 (23.9)	15 (24.1)	1.00 (Reference)	
GT	116 (51.7)	143 (53.3)	1.22 (0.73-1.99)	0.44	43 (44.7)	25 (40.3)	0.84 (0.35-1.99)	0.69
TT	63 (28.1)	79 (29.4)	1.35 (0.78-2.34)	0.27	30 (31.2)	22 (35.4)	1.30 (0.50-3.38)	0.58
GT+TT	179 (79.9)	222 (82.7)	1.26 (0.78-2.02)	0.33	73 (76.0)	47 (75.8)	0.99 (0.44-2.21)	0.9

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.7.5 Genetic variations in *XRCC6* and *XRCC7* genes and its relationship with clinicopathological parameters

The frequency distribution for the two polymorphic genes *XRCC6* and *XRCC7* according to the clinicopathological characteristics are detailed in table 5.28. Overall observations for the two SNPs suggested no significant correlation with either of the clinicopathological factors *i.e.* stage, tumor extension, lymph node involvement or metastasis.

On the basis of clinical stage, the patients were stratified into stage III and stage IV subjects. The genotypic distribution for each polymorphic site has been estimated and summarized as in table 5.28. In case of *XRCC6* polymorphic site $C^{61}G$ subjects with single allelic variant genotype (CG) were found to be represented more among stage IV (50.0%) subjects as compared to stage III patients (46.0%). However, the difference was not significant and a visible trend towards slight increased risk was observed for stage IV (OR=1.4, 95%C.I.=0.87-2.30, $p=0.15$). Similar to the above, subjects diagnosed with stage IV and carrying combined variant genotype ($CG+GG$) (59.8%) had a higher prevalence than stage III subjects (53.9%). Logistic regression analysis showed a trend towards association of the genotype for stage IV with heterozygous genotype (CG) (OR=1.49, 95%C.I.=0.93-2.3, $p=0.09$). The frequency of combined variant genotype ($GT+TT$) in case of *XRCC7* was almost similar among stage III (80.8%) and stage IV (80.9%) subgroup respectively. Taking wild type genotype as a reference, logistic regression analysis showed a trend of no correlation with combined variant genotype and stage IV tumor (OR=1.14, 95%C.I.=0.63-2.05, $p=0.64$). Hence, *XRCC7* lacked any relationship with clinical staging.

Primary tumor extension was next evaluated for the two polymorphic genes. The patients were classified into two subgroups namely, T3+T4 and Tx+T1+T2. None of the two polymorphic sites conferred any effect on primary tumor extension. In case of *XRCC6* SNP the lung cancer patients with mutant genotype (GG) were similar in frequency in both groups *i.e.* T3+T4 (8.8%) and Tx+T1+T2 (9.1%) respectively. For the polymorphic variant *XRCC7* the percentage of combined variant genotype ($CG+GG$) was similar among T3+T4 and Tx+T1+T2 (80.8% vs 80.9%) subgroups respectively. Odds ratio revealed no association of the genotype towards tumor size (OR=1.24, 95%C.I.=0.61-2.51, $p=0.54$).

In case of lymph node involvement the study group were stratified into two groups N0 and N1+N2+N3 *i.e.* those subjects without lymph node and those with lymph node involvement. In case of *XRCC6* no significant difference was observed in the frequency distribution of the combined variant genotype ($CG+GG$) among lung cancer patients

with N0 (53.8%) and N1+N2+N3 (56.4%) subgroups. Subjects with combined variant genotype ($GT+TT$) in case of *XRCC7* were almost similar in frequency for N0 and N1+N2+N3 groups (78.8% vs 80.9%). Hence, lymph node was observed not to play a role in modulating lung cancer. Further in case of distant metastasis, patients were divided into M0 and M1 subgroups *i.e.* those where tumor has not spread to other regions and those where tumor has spread to other regions. For the polymorphic variant of *XRCC6* double allelic variant genotype (GG) subjects were similar in frequency in both M0 and M1 groups (8.9% vs 9.1%) respectively. In case of *XRCC7* the frequency of subjects with mutant genotype (TT) was higher for subjects with distant metastasis 36.6% compared to subjects without distant metastasis 25.8%. Odds ratio revealed no association of mutant subjects with lung cancer risk. Therefore, it can be concluded the two polymorphic variants are not playing role in tumor progression as no significant association was observed for any clinicopathological factor.

Table 5.28: Relation between clinicopathological factors and polymorphic *XRCC6* and *XRCC7* (Primary Tumor Extension)

Genotype	Clinical Stage		AOR (95% CI) ^a	p ^b	Primary Tumor Extension		AOR (95%CI) ^a	p ^b
	III	IV			TX+ T1+ T2	T3+ T4		
XRCC6 61 C>G	178	142			68	241		
CC	82 (46.0)	57 (40.1)	1		27 (39.7)	109 (45.2)	1	
CG	82 (46.0)	71 (50.0)	1.4 (0.87-2.30)	0.15	35 (51.4)	110 (45.6)	0.85 (0.47-1.54)	0.61
GG	14 (7.8)	14 (9.8)	1.25 (0.81-1.93)	0.29	6 (8.8)	22 (9.1)	0.99 (0.59-1.67)	0.9
CG+GG	96 (53.9)	85 (59.8)	1.49 (0.93-2.3)	0.09	41 (60.2)	132 (54.7)	0.89 (0.51-1.55)	0.69
XRCC7 6721 G>T	178	142			68	241		
GG	34 (19.1)	27 (19.0)	1		13 (19.1)	46 (19.0)	1	
GT	96 (53.9)	65 (45.7)	1.02 (0.54-1.92)	0.93	38 (55.8)	118 (48.9)	1.19 (0.56-2.56)	0.63
TT	48 (26.9)	50 (35.2)	1.16 (0.83-1.60)	0.37	17 (25.0)	77 (31.9)	1.12 (0.75-1.69)	0.56
GT+TT	144 (80.8)	115 (80.9)	1.14 (0.63-2.05)	0.64	55 (80.8)	195 (80.9)	1.24 (0.61-2.51)	0.54

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

Table 5.29: Relation between clinicopathological factors and polymorphic *XRCC6* and *XRCC7* (Lymph Node and Metastasis)

Genotype	Lymph node Invasion		AOR (95% CI) ^a	p ^b	Metastasis		AOR (95%CI) ^a	p ^b
	N0	N1+ N2+ N3			No	Yes		
XRCC6 61 C>G	52	257			178	131		
CC	24 (46.1)	112 (43.5)	1		80 (44.9)	56 (42.7)	1	
CG	24 (46.1)	121 (47.0)	1.04 (0.55-2.95)	0.90	82 (46.0)	63 (48.0)	1.24 (0.76-2.0)	0.37
GG	4 (7.69)	24 (9.3)	1.18 (0.65-2.1)	0.57	16 (8.9)	12 (9.1)	1.09 (0.70-1.70)	0.68
CG+GG	28 (53.8)	145 (56.4)	1.36 (0.82-2.2)	0.22	98 (55.0)	75 (57.2)	1.26 (0.78-2.0)	0.33
XRCC7 6721 G>T	52	257			178	131		
GG	11 (21.1)	48 (18.6)	1		33 (18.5)	26(19.8)	1	
GT	22 (42.3)	134 (52.1)	1.32 (0.59-2.97)	0.48	99 (55.6)	57 (43.5)	0.88 (0.46-1.69)	0.70
TT	19 (36.5)	75 (29.1)	0.94 (0.62-1.43)	0.78	46 (25.8)	48 (36.6)	0.94 (0.62-1.43)	0.78
GT+TT	41 (78.8)	209 (81.3)	1.12 (0.52-2.3)	0.77	145 (81.4)	105 (80.1)	1.05 (0.58-1.92)	0.85

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

5.7.6 Modulation of lung cancer risk due to cumulative effect of *XRCC6* and *XRCC7* polymorphism

Possible cumulative risk of *XRCC6* 61 $C>G$ and *XRCC7* 6721 $G>T$ SNPs was also evaluated to find out any synergistic effect of both combined SNPs towards modulating lung cancer risk as shown in table 5.31 and henceforth we combined the adverse genotypes of *XRCC6* 61 $G>C$ and *XRCC7* 6721 $G>T$. It was observed that none of the genotypic variant for both the SNPs combined showed any effect towards modulating risk for lung cancer as odds ratio nearly equal to 1 was obtained. The frequency distribution for the combination of *XRCC6* and *XRCC7* suggested controls had 30.4% wild type subjects ($GG+GG$), 65.6% single allelic variant genotype/ heterozygous ($GC+GT$) and 4.0% mutant genotype ($CC+TT$). Whereas, lung cancer cases for the combination had 27.3% homozygous wild subjects ($GG+GG$), 67.5% heterozygous subjects ($GC+GT$) and 5.1% mutant subjects ($CC+TT$). Regression analysis suggested no association for any genotypic combination with lung cancer as shown in table 5.31. Further lung cancer risk was analyzed for the three histological subtypes. For subjects diagnosed with ADCC subtype and having single allelic genotype (GA) had a less frequency (53.4%) as compared to controls (65.5%). Logistic regression analysis showcased a trend of protective effect ($OR=0.5$, $95\%C.I.=0.25-1.13$, $p=0.1$) however it lacked significance. It was also observed that study subjects diagnosed with SCLC subtype revealed a high risk towards lung cancer for all the genotypic combinations. Subjects carrying single allele ($CG+GT$) of variant genotype showed a six-fold risk with approaching significance ($OR=6.7$, $95\%C.I.=0.84-53.8$, $p=0.07$). Similarly, combined variants for SCLC histology exhibited a seven fold risk towards lung cancer ($OR=7.0$, $95\%C.I.=0.88-56.02$, $p=0.06$).

Table 5.30: Genotypic distribution based on different genotypic combinations of *XRCC6* and *XRCC7* and their association with lung cancer risk assessment

XRCC6 61C>G	Controls N=125 (n%)	Overall		
		Cases N=117 (n%)	Overall AOR (95% CI) ^a	p ^b
CC/GG	38 (30.4)	32 (27.3)	1	
CG/GT	82 (65.6)	79 (67.5)	0.86 (0.47-1.57)	0.639
GG/TT	5 (4.0)	6 (5.1)	1.28 (0.33-4.87)	0.71
CG+GG/GT+TT	87 (69.6)	85 (72.6)	0.87 (0.48-1.58)	0.66

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

Table 5.31: Combinatorial assessment of XRCC6 and XRCC7 towards Lung cancer and Histology

XRCC6 61C>G	Controls N=125 (n%)	ADCC			SQCC			SCLC		
		Cases N=43 (n%)	Overall (95% CI) ^a	AOR p ^b	Cases N=49 (n%)	Overall (95% CI) ^a	AOR p ^b	Cases N=22 (n%)	Overall (95% CI) ^a	AOR p ^b
CC/GG	38 (30.4)	18 (41.8)	1		13 (26.5)	1		1 (4.5)	1	
CG/GT	82 (65.6)	23(53.4)	0.5 (0.25-1.13)	0.1	34 (69.3)	0.96 (0.4-2.16)	0.93	19 (86.3)	6.7 (0.84-53.8)	0.07
GG/TT	5 (4.0)	2 (4.6)	0.88 (0.14-5.35)	0.89	2(4.0)	1.11 (0.18-6.81)	0.90	2 (9.0)	13.9 (0.5-345.7)	0.10
CG+GG/GT+TT	87 (69.6)	25 (58.1)	0.55 (0.26-1.15)	0.11	36 (73.4)	0.9 (0.43-2.15)	0.94	21 (95.4)	7.0 (0.88-56.02)	0.06

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.8 Relationship of polymorphic variants of DSBRe gene *XRCC4* in lung cancer susceptibility

5.8.1 Demography of lung cancer subjects and controls with polymorphic *XRCC4*

Table 5.32: Demographic Characteristics among cases and controls of *XRCC4*

Variable	Cases N=346 (n%)	Controls N=367 (n%)	p ^a
Age(Years) Mean±SD Range	57.6±10.7	53.2±10.3	< 0.0001
Gender			
Male	298 (86.1)	304 (82.8)	0.0415
Female	48 (13.8)	63 (17.1)	
Smoking status			
Smokers	271(78.3)	257 (70.0)	0.0910
Non-smokers	75 (21.6)	110 (29.9)	
Pack Years Mean±SD	28.04±34.6	16.8±19.20	< 0.0001
Histology			
ADCC	117 (33.8)		
SQCC	147 (42.4)		
SCLC	79 (22.8)		
Others	3 (0.8)		
TNM staging			
I	3 (0.86)		
II	14 (4.0)		
III	160 (46.2)		
IV	145 (41.9)		
Unclassified	24 (6.9)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects,
(a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

The study characteristics for *XRCC4* SNP is enlisted in table 5.32. The study comprised 346 cases and 367 controls. The mean age for both the groups is illustrated in table 5.32 which was found to be 57.6±10.7 for cases and 53.2±10.3 for healthy controls enrolled in the study. The evaluation of the study based on gender suggested there was no difference in the frequency distribution of males and females between the two cohorts (p=0.0415)

86.1% (298) males and 13.8% (48) females for cases whereas there were 82.8% (304) males and 17.1% (63) females in controls. The smoking status calibrated the study into smokers and non-smokers. The patient group had slightly higher percentage of smokers as there were 78.3% (271) smokers and 21.6% (75) non-smokers. Controls of the study group had 70.0% (257) smokers and 29.9% (110) non-smokers. The difference was not observed to be significant. The smoking status was detailed further as mean pack years were also calculated for the both the study groups. The mean pack years was higher in cases compared to controls (28.04 ± 34.6 vs 16.8 ± 19.20). A significant difference ($p < 0.0001$) was observed between the pack years of cases and controls. Further considering the clinical parameters of the patients it was found that major percentage of subjects were diagnosed with SQCC (42.4%) histology, followed by ADCC (33.8%) and SCLC (22.8%). The distribution of patients on the basis of stage is also given in table 5.32. Stage III patients were highly represented with 46.2% and stage I were minimum with 0.86%. Stage II were 4.0% and 41.9% were stage IV subjects while remaining 6.6% were unclassified.

5.8.2 Allelic and genotypic distribution of *XRCC4* gene and its association towards lung cancer risk

The allelic and genotypic frequencies of polymorphic *XRCC4* is enlisted in table 5.33. Controls for the study were not following the hardy weinberg equilibrium ($\chi^2=9.4$, $p=0.002$) whereas cases were following HWE ($\chi^2=1.14$, $p=0.28$). The frequency distribution stated there was over representation of homozygous wild type genotype in controls (29.7%) compared to cases (25.4%). Whereas single allelic variant genotype was higher in cases (47.1%) as compared to controls (41.9%). There were 27.4% mutant subjects in cases compared to 28.3%. The difference of genotypic frequency between cases and controls was not significant ($\chi^2=2.2$, $p=0.31$). Regression analysis for the mutant subjects gave an OR of one fold (OR=1.07, 95% C.I.=0.86-1.31, $p=0.52$) as shown in table 5.33. Similarly, the combined variants (ID+DD) suggested no association with lung cancer (OR=1.23, 95% C.I.=0.87-1.72, $p=0.72$).

Table 5.33: Assessment of *XRCC4* intron 3 risk towards lung cancer susceptibility

XRCC4 Intron 3	Controls N=367 (n%)	Cases N=346 (n%)	Overall AOR (95% CI)^a	p^b
II	109 (29.7)	88 (25.4)	1	
ID	154 (41.9)	163 (47.1)	1.27 (0.88-1.84)	0.18
DD	104 (28.3)	95 (27.4)	1.07 (0.86-1.31)	0.52
ID+DD	258 (74.5)	258 (70.2)	1.23 (0.87-1.72)	0.72

to be cont'd on next page

Table 5.33: Assessment of *XRCC4* intron 3 risk towards lung cancer susceptibility (contd..)

XRCC4 Intron 3)	Controls N=367 (n%)	Cases N=346 (n%)	Overall AOR (95% CI) ^a	p ^b
I				
D				
MAF	0.50	0.49		

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.8.3 Genotypic and allelic distribution of *XRCC4* variant and risk towards lung cancer based on histology

As lung cancer is one of the most heterogenous solid tumor and the molecular pathology of each subtype is varied. Hence, the analysis for cancer predisposition was also carried out for different subtypes of lung cancer in relation to different genotypes of *XRCC4* and the data is summarized in table 5.34. The frequency distribution for ADCC subtype stated there were 29.9% wild subjects (*II*), 43.5% single allelic carriers *ID* and 26.5% mutant subjects *DD* in lung cancer cases. Controls were divided as 29.7% homozygous wild (*II*) subjects, 41.9% heterozygous subjects *ID* and 28.3% double allelic variants *DD*. Subjects with ADCC subtype and carrying heterozygous genotype (*ID*) were represented more as compared to controls (43.5 vs 41.9%). However, the difference between the frequency was not significant and a trend of no risk was observed (OR=1.06, 95%C.I.=0.64-1.77, p=0.80). On the same lines, subjects diagnosed with SQCC subtype and carrying single allelic variants (*ID*) were represented more in subjects as compared to controls (48.9 vs 41.9%). On comparing with wild subjects as odds ratio of 1.4 fold was obtained indicating a trend towards increased risk (OR=1.4, 95%C.I.=0.85-2.31, p=0.17). Similarly, subjects harboring combined variant genotype (*ID+DD*) exhibited an increased percentage 77.5% compared to 70.2% in controls. Risk analysis showcased a risk of 1.3 fold though the difference between the frequency distribution was not significant (OR=1.30, 95%C.I.=0.82-2.08, p=0.25) as shown in table 5.34.

Table 5.34: Polymorphic *XRCC4* intron 3 towards lung cancer based on Histology

Genes		ADCC			SQCC			SCLC		
XRCC4 Intron 3	Controls N=367 (n%)	Cases N=117 (n%)	AOR (95% CI)^a	p^b	Cases N=147 (n%)	AOR (95% CI)^a	p^b	Cases N=79 (n%)	AOR (95% CI)^a	p^b
II	109 (29.7)	35 (29.9)	1		33 (22.4)	1		20(25.3)	1	
ID	154 (41.9)	51 (43.5)	1.06 (0.64-1.77)	0.80	72 (48.9)	1.40 (0.85-2.31)	0.17	38(48.1)	1.21 (0.65-2.2)	0.53
DD	104 (28.3)	31 (26.4)	1.00 (0.74-1.34)	0.98	42 (28.5)	1.07 (0.81-1.43)	0.60	21(26.5)	1.09 (0.76-1.57)	0.61
ID+DD	258 (70.2)	103 (92.7)	1.04 (0.65-1.66)	0.85	114 (77.5)	1.30 (0.82-2.08)	0.25	59(74.6)	1.18 (0.66-2.1)	0.57

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.8.4 Genotypic distribution of *XRCC4* intron 3 variant based on smoking status and its association with risk of developing lung cancer

In order to evaluate the role of smoking in modulating the association between *XRCC4* gene polymorphism and lung cancer risk, the subjects were classified into smokers and non-smokers. Association analysis was carried out for the polymorphic site along with adjusted odds ratio with 95% C.I. within both the subgroups has been tabulated in table 5.35. Smokers possessing heterozygous genotype (ID) were over-represented in cases as compared to controls (48.3 % vs 45.1%). On comparing with wild type subjects no risk was observed towards lung cancer (OR=1.15, 95%C.I.=0.75-1.76, p=0.51) for smokers. Similarly, combined variants (ID+DD) showed no risk towards lung cancer development as shown in table 5.35.

Non-smokers harboring deletion at one allele (ID) had a high representation in controls (42.6%) compared to cases (45.1%). Taking wild type genotype as a reference a visible trend towards high risk was observed though no significance was obtained (OR=1.9, 95% C.I.=0.88-4.02, p=0.09).

Table 5.35: Genotypic distribution of *XRCC4* intron 3 based on smoking status and its association with risk of lung cancer

Genes	Smokers				Non Smokers			
XRCC4 Intron 3	Controls N=257 (n%)	Cases N=271 (n%)	AOR ^a (95%) CI	p^b	Controls N=110 (n%)	Cases N=75 (n%)	AOR ^a (95%) CI	p^b
II	68 (26.4)	70 (25.8)	1.0		41 (37.2)	18 (24.0)	1	
ID	116 (45.1)	131 (48.3)	1.15 (0.75-1.76)	0.51	38 (34.5)	32 (42.6)	1.9 (0.88-4.02)	0.09
DD	73 (28.4)	70 (25.8)	0.82 (0.63-1.07)	0.14	31 (28.1)	25 (33.3)	1.5 (0.99-2.29)	0.05
ID+DD	189 (73.5)	201 (74.1)	1.07 (0.72-1.60)	0.71	69 (62.7)	57 (76.0)	1.99 (1.01-3.9)	0.04

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

5.9 Survival analysis of BER genes *OGG1*, *MUTYH* and *XRCC1*

5.9.1 Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of polymorphic *OGG1* and *MUTYH* genes

The clinical characteristics and demography of the study undertaken for the overall survival of lung cancer patients in relation to the polymorphic *OGG1* and *MUTYH* is detailed in table 5.36. The study enrolled 251 study subjects with the mean age of the study being 58.5 ± 10.3 and range 28-83. There were 88.8% males and 11.1% females in the study. The study recorded maximum 82.4% smokers with remaining 17.5% non-smokers. Clinical factors such as histology, TNM staging, performance status and chemotherapeutic regimen are also listed in table 5.36. On the basis of histological classification there were 34.6% adenocarcinomas, 41.4% squamous cell carcinomas, 23.5% small cell carcinomas while there were 0.3% patients with unclassified histology. The TNM staging grouped the study subjects into 54.1% stage III patients and 44.6% stage IV patients and remaining 3.3% were unclassified. The study was analysed for the performance status ECOG and KPS. The classification according to performance status ECOG divided the study subjects into 45.1% subjects with ECOG score of 0-1, 38.6% with ECOG score of 2, 15.5% with ECOG score 3-4 and remaining 0.7 with unclassified ECOG score. Further, KPS score for the study classified the subjects into 64.1% subjects with KPS score between 100-80, 32.2% subjects with KPS score in range of 70-60, and 2.7% with KPS score of 50-40, and remaining with 0.7% with unclassified KPS score. There were total 172 subjects who received with chemotherapy data. 70 (40.6%) subjects out of total 172 received docetaxel along with cisplatin/carboplatin. 47 (27.3%) subjects were administered irinotecan in combination with cisplatin/carboplatin. Whereas, rest 55 (31.9%) were given pemetrexed together with cisplatin/carboplatin.

Table 5.36: Patients Characteristics and clinical features

Patients characteristics	No. of Patients, n(%) N=251
Age(Years)	
Mean $\pm$ SD Range	58.5 $\pm$ 10.3 (28-83)
Gender	
Male	223 (88.8)
Female	28 (11.1)

to be cont'd on next page

Table 5.36: Patients Characteristics and clinical features

Smoking status	
Smokers	207 (82.4)
Non-smokers	44 (17.5)
Histology	
ADCC	87 (34.6)
SQCC	104 (41.4)
SCLC	59 (23.5)
Others	1 (0.3)
TNM staging	
III	136 (54.1)
IV	112 (44.6)
Unclassified	3 (1.1)
ECOG performance status	
0-1	113 (45.0)
2	97 (38.6)
3-4	39 (15.5)
Others	2 (0.7)
KPS performance status	
100-80	161 (64.1)
70-60	81 (32.2)
50-40	7 (2.7)
Others	2 (0.7)
Median Follow up (months)	7.5
Events, Deaths	N = 219 (87.2%)
Chemotherapy Regimen	N = 172
Docetaxel + Cisplatin/Docetaxel + Carboplatin	70 (40.6)
Irinotecan + Cisplatin/Irinotecan + Carboplatin	47 (27.3)
Pemetrexed + Cisplatin/Pemetrexed + Carboplatin	55 (31.9)

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects, (a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

5.9.2 Association of polymorphic *OGG1*, *MUTYH* genetic variants with overall survival of lung cancer patients

Univariate and multivariate analysis for polymorphic *OGG1* and *MUTYH* snps is summarized in table 5.37. Univariate analysis is defined by Kaplan meir statistics which gives median survival time (MST) and survival graphs for each genotype. Log rank test performed in Kaplan meir statistics compares the MST for different genotypes consider-

ing the wild genotype as a reference. Kaplan meir statistics give an unadjusted hazard ratio whereas multivariate analysis using cox-regression provided an adjusted hazards ratio. Cox-regression model adjusted the hazards ratio for co-variables such as age, sex, smoking, histology, stage, KPS and ECOG. Overall the two SNPs *i.e.* *OGG1 Ser³²⁶Cys* and *MUTYH Gln³²⁴His* did not modulate the lung cancer prognosis. However, in case of *OGG1 Ser³²⁶Cys*, subjects carrying single allelic variant genotype (*CG*), there were 78.1% (25) subjects alive and 84.4% (185) subjects dead as shown in table 5.37. The median survival time was higher for the carriers of both wild double allelic genotype (*CC*) (MST=12.0) and homozygous variant genotype (*GG*) (MST=16.1, p=0.54) compared to heterozygous genotype (*CG*) (MST=7.2, p-value=0.42) though statistically insignificant (figure 5.17a). Further, multivariate analysis suggested that the HR values was less for patients with mutant genotype (H.R.=0.70, 95%C.I.=0.46-1.06, p=0.10) than for subjects with heterozygous subjects (H.R.=1.0, 95%C.I.=0.66-1.71, p=0.79) as shown in table 5.37.

In case of *MUTYH Gln³²⁴His* the percentage of alive and dead subjects was 40.6% and 51.1% respectively for heterozygous subjects (*Gln/His*). The difference in the MST for each genotypic combinations was not found to be significant although the median survival time for mutant (*CC*) genotype was almost double in comparison to the wild type (*GG*) genotype (MST=15.3 vs 7.6, Log rank p-value=0.61) (figure 5.17b). Similar to the above gene, hazard rate analysis suggested no association with lung cancer (OR=0.9, 95%C.I.=0.64-1.29, p=0.61) for mutant subjects as shown in table 5.37.

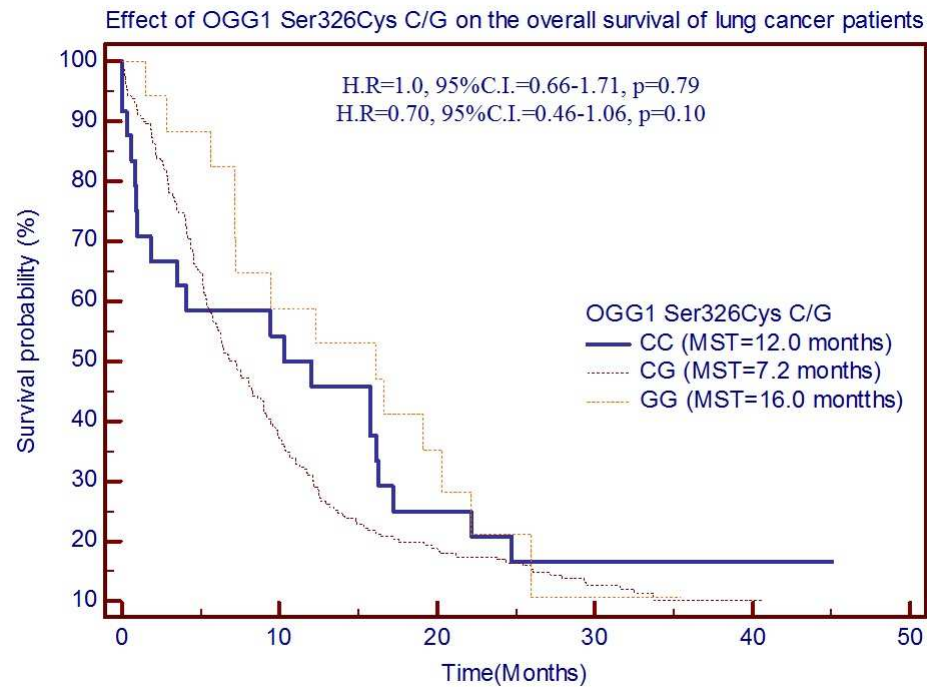
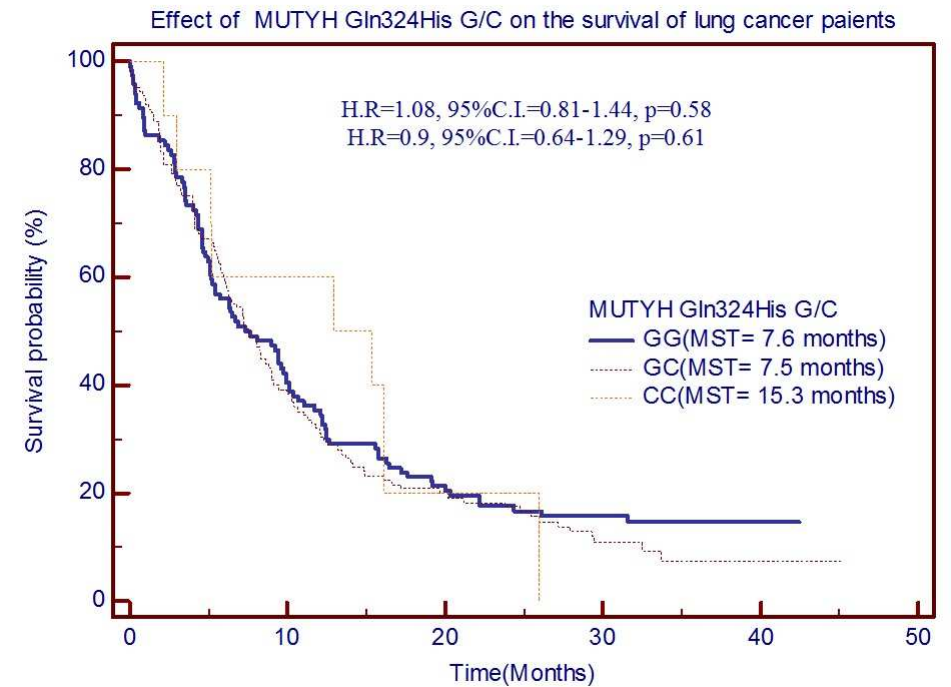
(a) *OGG1 Ser³²⁶Cys*(b) *MUTYH Gln³²⁴His*Figure 5.17: Kaplan Meier Overall Survival Curve for (a) *OGG1 Ser³²⁶Cys* (b) *MUTYH Gln³²⁴His*

Table 5.37: *OGG1* and *MUTYH* polymorphic variants and their association with overall survival

Genes	Cases N=251 (n%)	Death N=219 (n%)	Alive N=32 n(%)	Univariate analysis			Multivariate analysis	
				MST (months)	Log rank p-value	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	p-value
OGG1 Ser³²⁶Cys C/G								
CC	24 (9.5)	20 (9.1)	4 (12.5)	12.0		1	1	
CG	210 (83.6)	185 (84.4)	25 (78.1)	7.2	0.42	1.20 (0.78-1.85)	1.0 (0.66-1.71)	0.79
GG	17 (6.7)	14 (6.3)	3 (9.3)	16.1	0.54	0.81 (0.41-1.59)	0.70 (0.46-1.06)	0.10
CG+GG	227 (90.4)	199 (90.8)	28 (87.5)	7.3	0.50	1.16 (0.78-1.79)	0.98 (0.72-1.3)	0.92
MUTYH Gln³²⁴His G/C								
GG	116 (46.2)	98 (44.7)	18 (56.2)	7.6		1	1	
GC	125 (49.8)	112 (51.1)	13 (40.6)	7.5	0.50	1.09 (0.83-1.43)	1.08 (0.81-1.44)	0.58
CC	10 (3.9)	9 (4.1)	1 (3.1)	15.3	0.75	0.89 (0.46-1.72)	0.9 (0.64-1.29)	0.61
GC+CC	135 (53.7)	121 (55.2)	14 (43.7)	7.5	0.56	1.08 (0.82-1.41)	1.05 (0.79-1.39)	0.72

(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

5.9.3 Cox-regression analysis of the factors associated with overall survival of *OGG1* and *Mutyh* in lung cancer patients

Multivariate logistic regression analysis showed that age sex, smoking, histology, ECOG, KPS were not associated with the prognosis for mutant subjects in case of polymorphic *OGG1 Ser³²⁶Cys*. Stage for subjects carrying mutant genotype in case of *OGG1 Ser³²⁶Cys* was independent prognostic factor as shown in table 5.38.

Table 5.38: Results of cox-regression multivariate analysis for subjects with polymorphic *OGG1*

Variables	B	SE	Wald	P-value	Exp(b)	95%C.I.
OGG1 Ser ³²⁶ Cys	-0.351	0.214	2.682	0.101	0.703	0.462-1.069
AGE	0.034	0.023	2.1092	0.1464	1.0353	0.9882- 1.084
SEX	1.481	0.800	3.423	0.064	0.227	0.047 - 1.083
Smoking	-0.646	0.643	1.009	0.314	0.523	0.149 - 1.836
Histology	-0.039	0.302	0.017	0.895	0.961	0.533 - 1.732
Stage	0.778	0.245	10.024	0.001	2.178	1.348 to 3.518
ECOG	-0.096	0.361	0.070	0.789	0.908	0.449 to 1.837
KPS	-0.234	0.416	0.316	0.573	0.791	0.351 to 1.781

B-Regression coefficient; SE-Standard error; Wald-Wald value;HR-Hazard ratio; 95%C.I.-confidence intervals

5.9.4 Association of polymorphic *OGG1*, *MUTYH* and overall survival according to histological subtypes of lung cancer patients

The two polymorphic variant were evaluated for their potential to act as the predictor of survival of lung cancer patients based on histology. The study was stratified according to histology as tabulated in tables 5.39, 5.40, 5.41. Data on association of *OGG1* and *MUTYH* with survival of lung cancer subjects was analyzed and summarized in tables below. Our data observed that in case of *OGG1 Ser³²⁶Cys* the proportion of subjects having heterozygous genotype (*CG*) was 49.3% in ADCC, 81.7% in SQCC and 83.0% in SCLC subjects. Overall survival analysis revealed that subjects with ADCC subtype and carrying mutant genotype (*GG*) as shown in table 5.39 had almost double survival time as compared to ADCC subjects with heterozygous (*CG*) genotype (MST=13.0, Log

rank p-value=0.49 vs 7.6, log rank p-value=0.84). As shown in table 5.40 SQCC patients carrying double allelic variant genotype (*GG*) had a better MST of 14.2 months compared to 10.2 months for subjects carrying both allele for homozygous wild type genotype (*CC*) (log rank p-value=0.78). Similarly, subjects diagnosed with SCLC subtype and carrying heterozygous genotype (*CG*) exhibited a higher MST compared to patients carrying homozygous wild genotype (*CC*) (20.3 vs 12.0, log rank p-value=0.85). Subjects harboring combined variant genotype (*CG+GG*) for SCLC subtype had a frequency of 91.5%. Cox-regression analysis showcased a relatively high death rate with shorter survival time (MST=8.2 vs 10.0: H.R.=1.45, 95%C.I.=0.47-4.42, p=0.51) as compared to SCLC subjects with wild genotype.

MUTYH Gln³²⁴His was found to play no role in the prognosis of lung cancer subjects for any histological subtype. The proportion of subjects carrying *GC+CC* genotype was 44.3% in ADCC subtype, 59.6% in SQCC subtype and 57.6% under SCLC subtype as shown in tables 5.39, 5.40 and 5.41 respectively. Kaplan Meir analysis estimated the MST for each category. Lung cancer subjects with ADCC subtype and harbouring wild type genotype (*GG*) showed the highest MST of 9.2 months as compared to 7.1 months for single allelic carriers (*GC*) and 5.1 months for ADCC subjects with mutant genotype (*CC*). In case of SQCC patients, a higher MST was observed for mutant genotype (*His/His*) as compared to the wild (*GG*) subjects (10.6 vs 8.0, log rank p-value=0.7). Similarly, subjects with SCLC subtype and carrying double allelic variant genotype (*CC*) showed a better MST compared to other two genotypes (*GG*= 3.5 vs 15.3, log rank p-value=0.51, *GC* =3.5 vs 9.0, log rank p-value=0.88).

Table 5.39: Association of *OGG1* and *MUTYH* SNPs with survival of lung cancer patients according to Tumor histology ADCC

Genes	ADCC					
OGG1 Ser ³²⁶ Cys	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=88	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	6 (6.8)	8.3		1	1	
CG	76 (86.3)	7.6	0.84	1.09 (0.45-2.62)	0.90 (0.32-2.5)	0.85
GG	6 (6.8)	13.0	0.49	0.64 (0.17-2.39)	-	-
CG+GG	81 (92.0)	8.3	0.9	1.05 (0.43-2.55)	0.7 (0.28-2.1)	0.64
MUTYH Gln ³²⁴ His	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=88	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	49 (55.6)	9.2		1	1	
GC	36 (40.9)	7.1	0.68	1.1 (0.68-1.76)	1.16 (0.68-1.95)	0.55
CC	3 (3.4)	5.1	0.84	0.86 (0.22-3.29)	0.79 (0.36-1.76)	0.58
GC+CC	39 (44.3)	8.3	0.77	1.06 (0.68-1.64)	1.03 (0.65-1.65)	0.87

to be cont'd on next page

Table 5.39: Association of *OGG1* and *MUTYH* SNPs with survival of lung cancer patients according to Tumor histology ADCC (contd..)

Genes	ADCC
(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG	

Table 5.40: Association of *OGG1* and *MUTYH* SNPs with survival of lung cancer patients according to Tumor histology SQCC

Genes	SQCC					
OGG1 Ser ³²⁶ Cys	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=104	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	13 (12.5)	10.2		1	1	
CG	85 (81.7)	6.4	0.78	1.09 (0.59-2.01)	1.24 (0.61-2.49)	0.54
GG	6 (5.7)	14.2	0.78	-	-	-
CG+GG	91 (87.5)	6.8	0.8	1.07 (0.57-1.98)	1.19 (0.59-2.39)	0.61
MUTYH Gln ³²⁴ His	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=104	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	42 (40.3)	8.0		1	1	
GC	58 (55.7)	6.4	0.44	1.18 (0.77-1.82)	0.89 (0.50-1.58)	0.69
CC	4 (3.8)	10.6	0.7	1.16 (0.38-3.50)	-	-
GC+CC	62 (59.6)	7.3	0.48	-	1.17 (0.76-1.80)	0.46

(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.41: Association of *OGG1* and *MUTYH* SNPs with survival of lung cancer patients according to Tumor histology SCLC

Genes	SCLC					
OGG1 Ser ³²⁶ Cys	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=59	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	5 (8.4)	12.0		1	1	
CG	49 (83.0)	7.3	0.28	1.70 (0.74-3.8)	1.7 (0.57-5.4)	0.31
GG	5 (8.4)	20.3	0.85	0.88 (0.23-3.3)	-	-
CG+GG	54 (91.5)	8.2	0.33	1.62 (0.70-3.74)	1.45 (0.47-4.42)	0.51
MUTYH Gln ³²⁴ His	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=59	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	25 (42.3)	3.5		1	1	
GC	31 (52.2)	9.0	0.88	1.04 (0.60-1.79)	0.90 (0.47-1.70)	0.75
CC	3 (5.0)	15.3	0.51	0.67 (0.23-1.94)	0.84 (0.41-1.69)	0.62
GC+CC	34 (57.6)	10.4	0.89	0.96 (0.57-1.63)	0.77 (0.42-1.30)	0.39

to be cont'd on next page

Table 5.41: Association of *OGG1* and *MUTYH* SNPs with survival of lung cancer patients according to Tumor histology SCLC (contd..)

Genes	SCLC
(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG	

5.9.5 Association of *OGG1 Ser³²⁶Cys* and *MUTYH Gln³²⁴His* polymorphism with treatment response towards chemotherapeutic treatment and their correlation with objective response

The treatment outcome for the two SNPs was analyzed making use of response factors. According to the RECIST criteria the patients were divided into good and bad responders. Patients with complete remission (CR) and partial remission (PR) together were defined as good responders and patients with stable disease (SD) and partial disease (PD) were classified as poor responders. Adjusted odds ratio was calculated for each polymorphic site taking wild type genotype as a reference and adjusting for factors such as age, gender, smoking, stage, histology, KPS, ECOG, and chemotherapeutic regimens. There were in total 148 patients with objective response outcome data. None of the polymorphic sites of *OGG1* and *MUTYH* revealed any significant difference in the frequency distribution of genotype. *OGG1 Ser³²⁶Cys* was observed to be classified as good responders irrespective of the genotype as shown in the table 5.42. Good responders with heterozygous genotype (*CG*) (85.8%) had a higher representation of frequency compared to bad responders (80.9%) with the above genotype. An adjusted odds ratio of nearly 2.3 fold was observed (OR=2.32, 95%C.I.=0.61-8.76, p=0.21) for the same. The representation of double allelic variant genotype (*GG*) was almost similar in both good and bad responders (5.8 vs 6.3). Hence no significant difference was observed however, a visible trend classifying mutant genotype (*GG*) as a good responders was observed (OR=1.92, 95%C.I.=0.33-11.02, p=0.46).

In case of *MUTYH Gln³²⁴His* the frequency of heterozygous genotype was slightly higher among good responders as compared to bad responders (51.7% vs 46.0%) respectively. Subjects with combined variant genotype (*GC+CC*) showed a higher percentage among good responders (55.2%) compared to bad responders (47.2%) as shown in table 5.42.

Table 5.42: Effect of *OGG1* and *MUTYH* SNPs on treatment outcome

Genotype	Good response CR+PR n(%) N = 85	Poor response SD+PD n(%) N = 63	AOR (95%C.I.) ^a	p ^b
OGG1 Ser³²⁶Cys C/G				
CC	7 (8.2)	8 (12.6)	1	
CG	73 (85.8)	51 (80.9)	2.32 (0.61-8.76)	0.21
GG	5 (5.8)	4 (6.3)	1.92 (0.33-11.02)	0.46
CG+GG	78 (91.6)	55 (87.3)	1.77 (0.54-5.78)	0.39
MUTYH Gln³²⁴His G/C				
GG	38 (44.7)	33 (52.3)	1	
GC	44 (51.7)	29 (46.0)	1.10 (0.50-2.40)	0.79
CC	3 (3.5)	1 (1.5)	0.93 (0.25-3.45)	0.91
GC+CC	47 (55.2)	30 (47.6)	1.08 (0.50-2.32)	0.82

(a)-Adjusted Odds ratio, (b)- 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, histology, stage, KPS, ECOG. Abbreviations: CR-Complete reponse, PR-Partial reponse, SD-Stable disease, PD-Progressive disease, AOR-Adjusted odds ratio, C.I.-Confidence interval

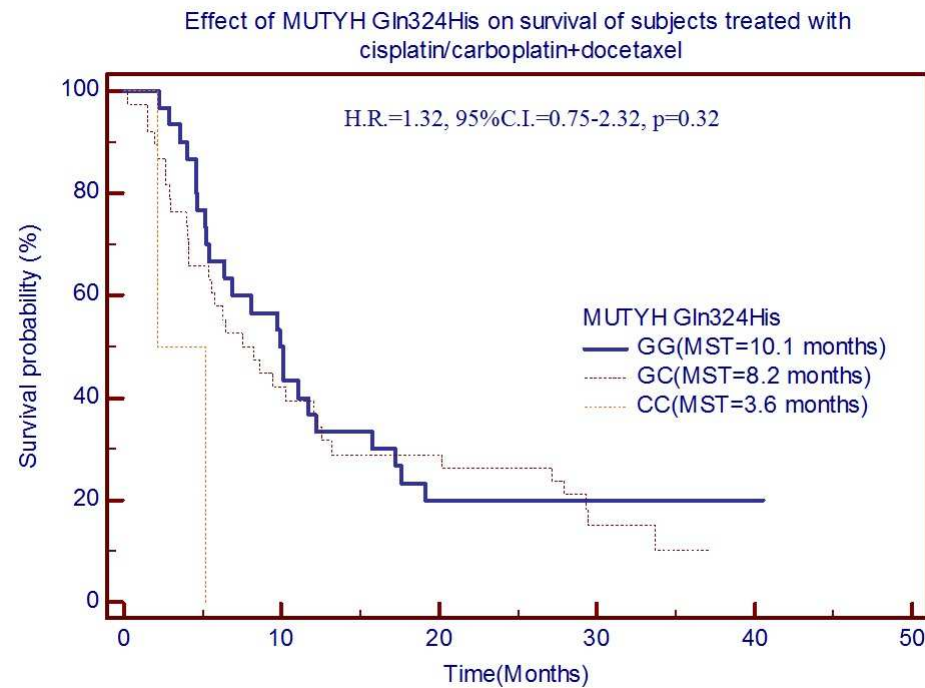
5.9.6 Overall survival of lung cancer patients according to different chemotherapeutic regimens in relevance to polymorphic *OGG1* and *MUTYH* genes

Lung cancer subjects were administered chemotherapeutic drugs in a cyclic manner. Patients were analyzed for independent effect of drugs on survival. Lung cancer subjects were given either of the three regimens based on the cisplatin/ carboplatin + docetaxel, or cisplatin/ carboplatin + pemetrexed, or cisplatin/ carboplatin + irinotecan. In case of *OGG1 Ser³²⁶Cys*, subjects administered cisplatin/ carboplatin + docetaxel, the frequency distribution showed that there were 87.1% of single allelic carriers (*CG*) and 5.7% of double allelic carriers (*GG*) as shown in table 5.43. The MST for double allelic carriers (*GG*) was observed to be 15.7 months which was the highest as compared to the MST of wild type genotype. However, the log rank test did not observe any significant difference in the MST for the three genotypic combinations as shown in table 5.43. The cox-regression analysis for single allelic carriers did not show any association with survival (MST=1.27, 95%C.I.=0.31-5.15, p=0.7) influence on hazard rate. 2.8% of subjects were observed to be of mutant genotype (*GC*) and 42.8% were wild type subjects (*GG*) for *MUTYH Gln³²⁴His*. MST for subjects with both allele of mutant genotype (*CC*) was

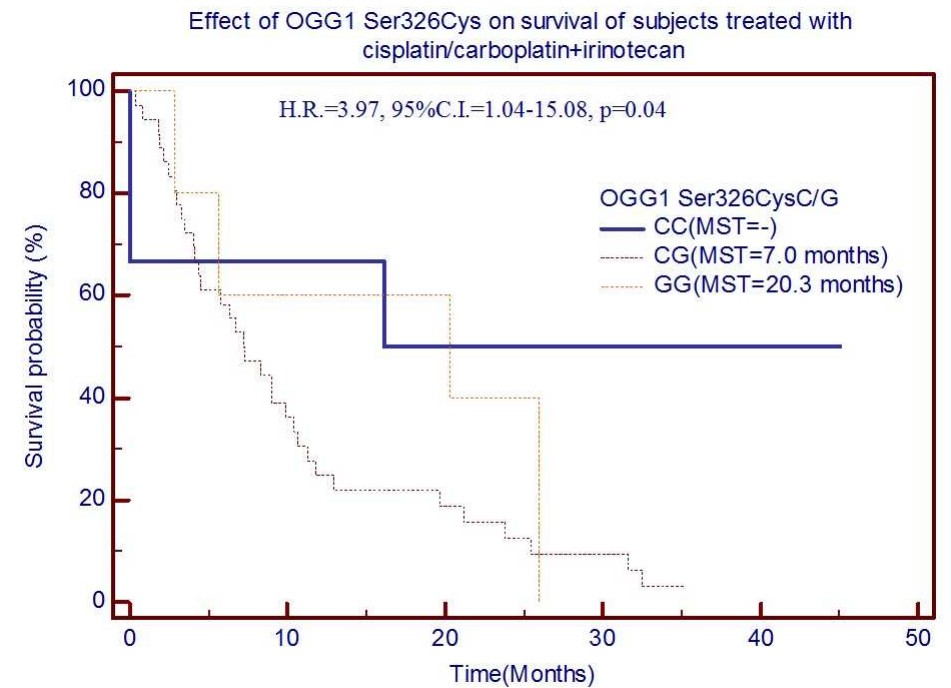
observed to be the smallest and significantly different when compared to subjects with wild genotype (*GG*) (3.6 vs 10.1, Log rank $p = \mathbf{0.0096}$) with a univariate death rate of 5.34 fold as shown in table 5.43 (figure 5.18a).

Irinotecan was given majorally to subjects with SCLC subtype along with cisplatin / carboplatin and it was observed that carriers of single allelic variant (*CG*) genotype showed a significant survival of 7.0 months (Log rank p -value = $\mathbf{0.04}$) compared to 20.3 months for mutant subjects (*GG*) as shown in figure 5.18b. A high hazard rate of three-fold was observed for univariate analysis when adjusted for co-variables a high death rate was sequentially observed for heterozygous (*CG*) genotype (H.R.=3.97, 95%C.I.=1.04-15.08, $p=\mathbf{0.04}$). In case of *MUTYH Gln³²⁴His* 10.4 months was observed to be the MST for double allelic variant genotype (*GC+CC*) as tabulated in table 5.44 (figure 5.19b).

The third group of subjects was administered Cisplatin/carboplatin + pemetrexed. Subjects with wild type genotype (*CC*) showed a poor MST of 0.56 months as compared to 9.1 months in case of heterozygous subjects (*CG*) which was significantly different (Log rank p value = $\mathbf{<0.0001}$) as shown in figure 5.19a. The hazards rate consecutively was very small for the same genotype (HR=0.051, 95%C.I.=0.0076-0.349, $p=\mathbf{0.0025}$). MST improved further for the double allelic variant carriers (*GG*) (MST =16.5, $p=\mathbf{0.038}$). As the sample size was very small for double allelic carriers the Cox regression could not solve the equation. Furthermore, for *MUTYH Gln³²⁴His* the survival analysis for the chemotherapeutic drugs independently did not show any significant alteration. However, there were 63.3% wild type subjects (*GG*), 32.7% single allelic carriers (*GC*) and 3.6% double allelic carriers (*CC*) as indicated in table 5.45.

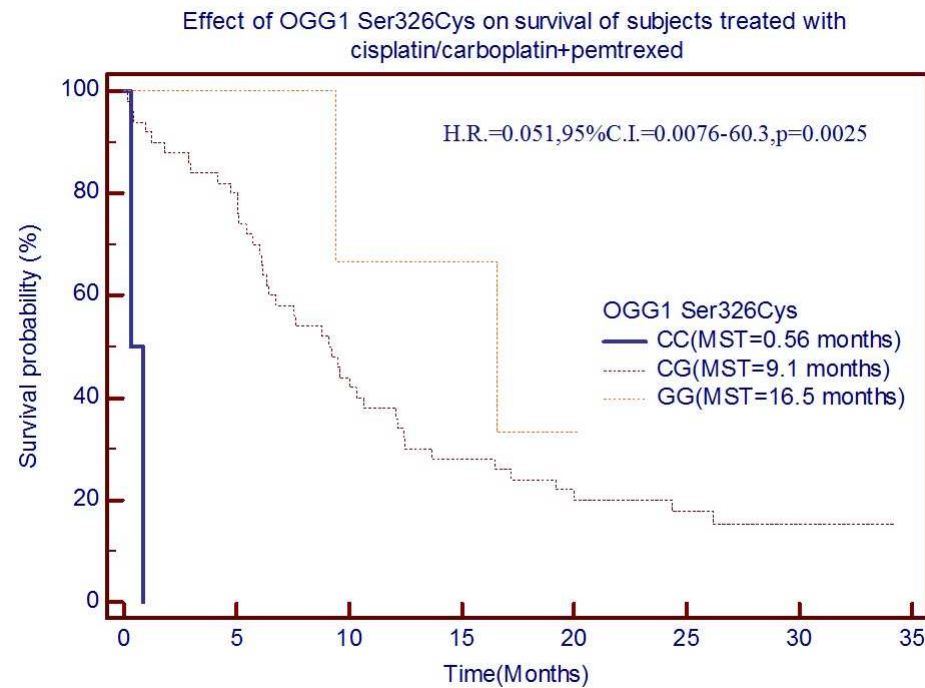


(a) Cisplatin/Carboplatin + Docetaxel

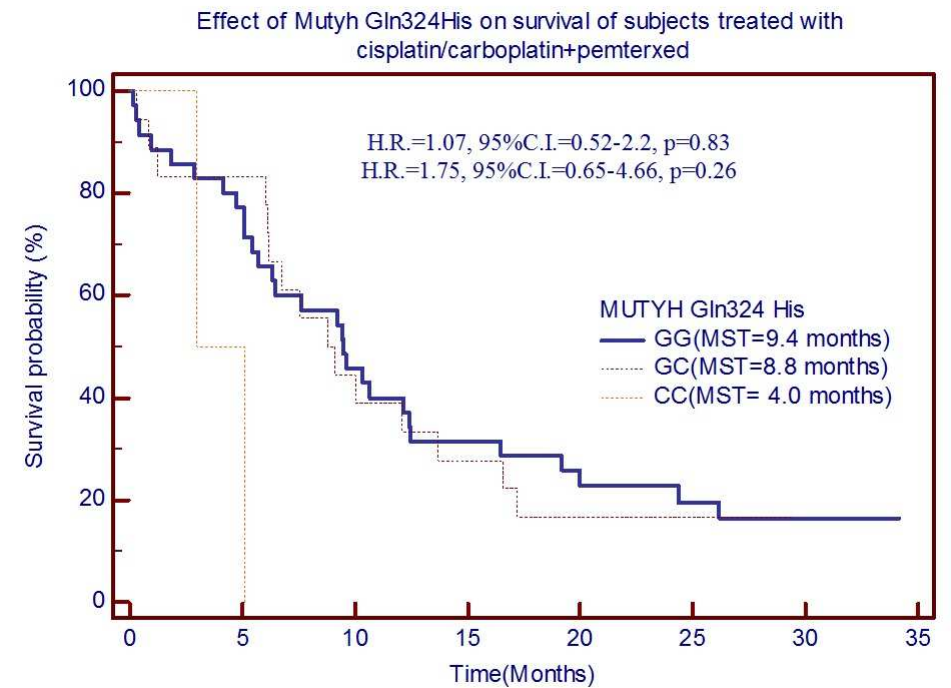


(b) Cisplatin/Carboplatin + Irinotecan

Figure 5.18: Kaplan Meier Overall Survival Curve for *MUTYH* Gln³²⁴His treated with (a) Cisplatin/Carboplatin + Docetaxel, for *OGG1* Ser³²⁶Cys treated with (b) Cisplatin/Carboplatin + Irinotecan



(a) Cisplatin/Carboplatin + Pemetrexed



(b) Cisplatin/Carboplatin + Pemetrexed

Figure 5.19: Kaplan Meier Overall Survival Curve for *OGG1 Ser³²⁶Cys* treated with (a) Cisplatin/Carboplatin + Pemetrexed for *OGG1 Ser³²⁶Cys* treated with (b) Cisplatin/Carboplatin + Pemetrexed

Table 5.43: Association of *OGG1* and *MUTYH* SNPs with chemotherapy regimen (Cisplatin/Carboplatin+Docetaxel) and Overall Survival

Genes	Cisplatin/Carboplatin+Docetaxel					
OGG1 Ser ³²⁶ Cys C/G	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=70	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	5 (7.1)	12.1		1	1	
CG	61 (87.1)	7.5	0.9	0.94	1.27 (0.31-5.15)	0.7
GG	4 (5.7)	15.7	0.24	0.46	-	-
CG+GG	65 (92.8)	8.0	0.82	0.90	1.24 (0.31-4.86)	0.75
MUTYH Gln ³²⁴ His G/C	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=70	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	30 (42.8)	10.1		1	1	
GC	38 (54.2)	8.2	0.50	1.19 (0.71-2.01)	1.32 (0.75-2.32)	0.32
CC	2 (2.8)	3.6	0.0096	5.34 (0.23-21.9)	-	-
GC+CC	40 (57.1)	6.4	0.39	1.25 (0.35-2.08)	1.37 (0.79-2.38)	0.26

(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.44: Association of *OGG1* and *MUTYH* SNPs with chemotherapy regimen (Cisplatin/Carboplatin+Irinotecan) and Overall Survival

Genes	Cisplatin/Carboplatin+Irinotecan					
OGG1 Ser ³²⁶ Cys C/G	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=47	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	6 (12.7)	-		1	1	
CG	36 (76.5)	7.0	0.04	2.96 (1.34-6.57)	3.97 (1.04-15.08)	0.04
GG	5 (10.6)	20.3	0.53	1.59 (0.36-7.50)	-	-
CG+GG	41 (87.2)	7.3	0.05	2.89 (1.31-6.3)	3.12 (0.81-12.0)	0.09
MUTYH Gln ³²⁴ His G/C	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=47	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	17 (36.1)	6.7		1	1	
GC	28 (59.5)	9.0	0.86	1.05 (0.55-2.01)	0.93 (0.40-2.14)	0.87
CC	2 (4.2)	-	0.3	0.36 (0.09-1.44)	0.74 (0.12-4.32)	0.74
GC+CC	30 (63.8)	10.4	0.98	0.99 (0.52-1.89)	0.90 (0.39-2.08)	0.81

(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.45: Association of OGG1 and MUTYH SNPs with Chemotherapy Regimen (Cisplatin/Carboplatin+Pemetrexed) and Overall Survival

Genes	Cisplatin/Carboplatin+Pemetrexed					
OGG1 Ser ³²⁶ Cys C/G	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=55	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	2 (3.6)	0.56		1	1	
CG	50 (90.9)	9.1	< 0.0001	0.076 (0.0005-10.8)	0.051 (0.0076-60.34)	0.0025
GG	3 (5.4)	16.5	0.038	0.194 (0.002-1.38)	-	-
CG+GG	53 (96.3)	9.4	< 0.0001	0.07 (0.0004-11.4)	0.056 (0.0076-60.33)	0.0021
MUTYH Gln ³²⁴ His G/C	Cases n(%)	Univariate analysis			Multivariate analysis	
	N=55	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	35 (63.6)	9.4		1	1	
GC	18 (32.7)	8.8	0.8	1.07 (0.57-2.01)	1.07 (0.52-2.2)	0.83
CC	2 (3.6)	4.0	0.05	3.54 (0.27-46.5)	1.75 (0.65-4.66)	0.26
GC+CC	20 (36.3)	7.5	0.60	1.16 (0.63-2.15)	1.16 (0.57-2.33)	0.67

(a)-Unadjusted Hazards ratio for Kaplan Meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

5.10 Survival analysis of BER gene *XRCC1*

5.10.1 Clinical characteristics of patients with polymorphic *XRCC1* gene

The clinical characteristics of the lung cancer patients for undertaken for evaluating survival analysis is detailed in table 5.46. The study reported 241 lung cancer subjects with an average median follow up time of 6.5 months. There were 193 (80.08%) events or deaths reported in the study. The mean age of patients group was 58.71±10.38 years. There were 215 (89.2%) males and 26 (10.7%) females that comprised the study group. Out of 241 patients 200 (82.9%) subjects were smokers and 41 (17.0%) were non-smokers. Clinical parameters such as histology, TNM staging, performance status and chemotherapeutic regimens are also enumerated in table 5.46. Pathological classification revealed there were 83 (34.4%) adenocarcinoma, 101 (41.9%) squamous cell carcinoma and 55 (22.8%) small cell carcinoma patients. The TNM staging categorization revealed 130 (53.9%) cases were of stage III and 111 (46.0%) were stage IV patients. Likewise, the lung cancer patients were also distributed accordingly to performance status. Eastern Cooperative

Oncology Group (ECOG) performance status (PS) and Karnofsky's performance scores (KPS) are two important criteria's of performance of lung cancer subjects treated with platinum based chemotherapy. In our study there were 127 (52.6%) patients with ECOG PS of [0-1], 73 (30.2%) were having ECOG PS of [2] and 41 (17.0%) subjects were ECOG PS [3] patients. Based upon Karnofsky's performance scores (KPS), there were 60.9% (147) patients with KPS score between [100-80], 30.7% (74) with KPS score of [70-60] and 8.4% (20) with KPS score between [50-40]. Further, out of 241 patients with follow up data, 165 were given platinum based doublet chemotherapy as first-line treatment. Of these 165 patients, 65 (39.3%) received cisplatin/carboplatin drug in combination with docetaxel, 44 (26.6%) received cisplatin/carboplatin in combination with irinotecan and 56 (33.9%) received cisplatin/carboplatin combination with pemetrexed.

Table 5.46: Patients Characteristics and clinical features of *XRCC1*

Patients Characteristics	No.of patients,(n%) N=241
Age (Years) Median±SD Range	58.7±10.38
Gender Male Female	215 (89.2) 26 (10.7)
Smoking status Smokers Non-smokers	200 (82.9) 41 (17.0)
Histology ADCC SQCC SCLC Others	83 (34.4) 101 (41.9) 55 (22.8) 2 (0.8)
TNM staging III IV	130 (53.9) 111 (46.0)
ECOG performance status 0-1 2 3-4	127 (52.6) 73 (30.2) 41 (17.0)
KPS performance status 100-80 70-60 50-40	147 (60.9) 74 (30.7) 20 (8.4)
Median follow up (months)	6.56
Events, Deaths	193 (80.8)
Chemotherapy Regimen	N= 165

to be cont'd on next page

Table 5.46: Demographic Characteristics among cases and controls of *XRCC1* (contd..)

Docetaxel +Cisplatin /Docetaxel +Carboplatin	65 (39.3)
Irinotecan+ cisplatin/ Irinotecan +Carboplatin	44 (26.6)
Pemetrexed+Cisplatin/Pemetrexed +Carboplatin	56 (33.9)

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects, (a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

5.10.2 Allelic and genotypic distribution of *XRCC1* genetic variants and their association with overall survival of lung cancer patients

Lung cancer subjects were assessed with relevance to polymorphic *XRCC1* for their roles in risk analysis in the prior sections. In the present work polymorphic *XRCC1* was also analyzed for its influence on survival. Table 5.47 shows the five polymorphic variants of *XRCC1* i.e. *XRCC1 Arg³⁹⁹Gln*, *Arg¹⁹⁴Trp*, *Pro²⁰⁶Pro*, *Gln⁶³²Gln*, and *Arg²⁸⁰His* and their association with overall survival. MST was calculated for each genotypic group and compared with reference group using Kaplan Meir analysis. Further, Cox regression analysis was used to evaluate the adjusted hazard ratio. Our analysis revealed that there were a total of 241 study subjects, out of which 193 subjects were dead during follow up period and 48 were alive.

Our data in the present study showed that the genotypic distribution of *XRCC1 Arg³⁹⁹Gln* ($\chi^2=1.9$, df=2; p=0.15), *Arg¹⁹⁴Trp* ($\chi^2=1.45$, df=2; p=0.22), *Pro²⁰⁶Pro* ($\chi^2=2.8$, df=2; p=0.09), *Gln⁶³²Gln* ($\chi^2=1.26$, df=2; p=0.26) did not deviate from Hardy Weinberg equilibrium. On the contrary *XRCC1 Arg²⁸⁰His* ($\chi^2=50.3$, p=**0.0001**) was not obeying the Hardy Weinberg equality law.

As shown in table 5.47 in case of *XRCC1 Arg³⁹⁹Gln* the frequency distribution suggested there were 58.3% subjects alive while 52.8% study subjects were dead carrying single allelic genotype (*GA*). While in subjects carrying double allelic variant genotype (*AA*), there were 20.8% subjects alive with 17.0% study subjects dead. A stepwise increase in median survival time (MST) with increasing number of variant allele was observed. Individuals with wild type genotype (*GG*) had an MST of 6.3 months, those with heterozygous genotype (*GA*) had an MST of 6.4 months (Log rank p-value=0.93), and mutant genotype (*AA*) had the highest MST of 9.4 months (Log rank p-value=0.43) as shown

in figure 5.20a. Accordingly applying the cox-proportional hazards model after adjusting for age, sex, smoking, stage, ECOG, KPS and histology, our data showed that subjects carrying two copies of the variant allele (*AA*) showed a trend of decreased death rate (H.R=0.87, 95%C.I.=0.56-1.35, p=0.54) as compared to individuals having the wild type genotype (*GG*). Whereas combined variant genotype (*GA+AA*) showed no association with death rate (HR=0.98, 95%C.I.=0.72-1.34, p=0.92) as shown in table 5.47.

Further in case of *XRCC1 Arg¹⁹⁴ Trp* the frequency distribution revealed that among all the patients carrying single copy of variant allele *GA* there were 22.9% subjects who were alive while 20.2% subjects dead. Survival analysis further in *XRCC1 Arg¹⁹⁴ Trp* revealed that wild type (*CC*) genotype showed an increased OS (MST=7.2) in comparison to lung cancer subjects who were carriers of single allelic variant (*CT*) (MST=5.9, Log rank p value=0.97) as figure 5.20b depicts.

Study population for *XRCC1 Pro²⁰⁶ Pro* showed the percentage of subjects who were carrying single copy of variant genotype (*AG*) and who were alive was 45.8% while 51.2% were dead. On the other hand there were 10.4% subjects who were alive and 9.3% subjects were dead with mutant genotype (*GG*) as shown in table 5.47. Both single allelic genotype (*GA*) (HR=1.02, 95%C.I.=0.75-1.39, p=0.87) and double allelic variants (*GG*) (HR=1.15, 95%C.I.=0.68-1.94, p=0.60) did not show any association with hazard rate (figure 5.20c).

Furthermore, the next SNP analyzed was *XRCC1 Gln⁶³² Gln*. In line with the analysis of *XRCC1 Arg³⁹⁹ Gln* it was observed that patients carrying the double allelic variants (*AA*) for *Gln⁶³² Gln* had a better survival rate as MST for mutant genotype (*AA*) was higher in comparison to the other genotypes (*AA*= 8.0, *GG*=6.8, *GA*=6.4 Log rank p-value=0.64) (figure 5.21a). After adjusting with Cox regression analysis death rate did not show any significant association with any genotype however, heterozygous genotype (*GA*) revealed an H.R of 1.21 (95%C.I.=0.80-1.84, p=0.35).

XRCC1 Arg²⁸⁰ His polymorphism was also analyzed for its association with overall survival and unlike other SNPs it was observed that homozygous wild type (*GG*) genotype had the highest MST of 12.0 months whereas carriers of single allelic variants (*GA*) and double allelic variants (*AA*) had an MST of 6.8 (Log rank p-value=0.22) and 6.3 (Log rank p-value=0.98) months respectively (figure 5.21b). Cox proportional hazard model, on evaluation suggested heterozygous genotype (*GA*) depicted a non-significant risk for mortality (H.R=1.32, 95%C.I.=0.825-2.14, p=0.24).

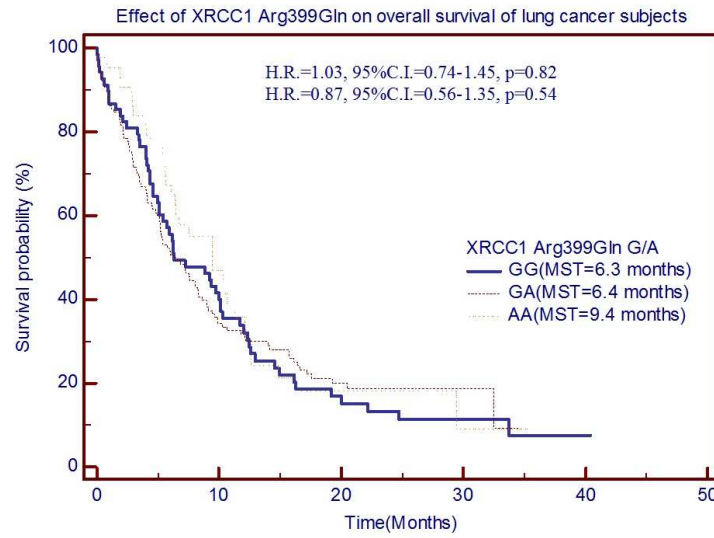
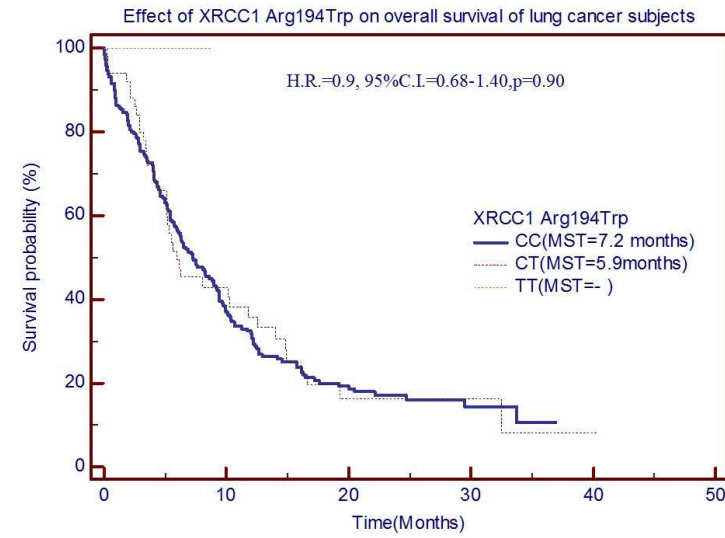
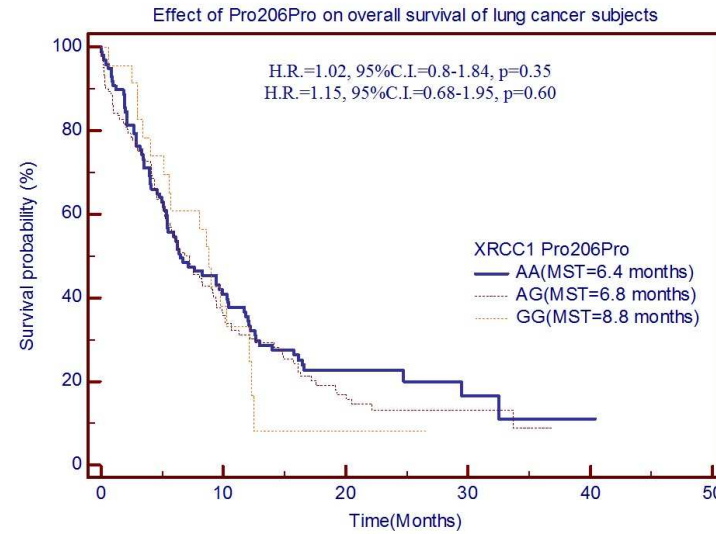
(a) *XRCC1 Arg³⁹⁹Gln*(b) *XRCC1 Arg¹⁹⁴Trp*(c) *XRCC1 Pro²⁰⁶Pro*

Figure 5.20: Kaplan Meier Overall Survival Curve for (a) *XRCC1 Arg³⁹⁹Gln* (GG:reference), (b) *XRCC1 Arg¹⁹⁴Trp* (CC:reference), (c) *XRCC1 Pro²⁰⁶Pro* (AA:reference) for lung cancer patients. H.R.:Hazard ration, MST:Median survival time

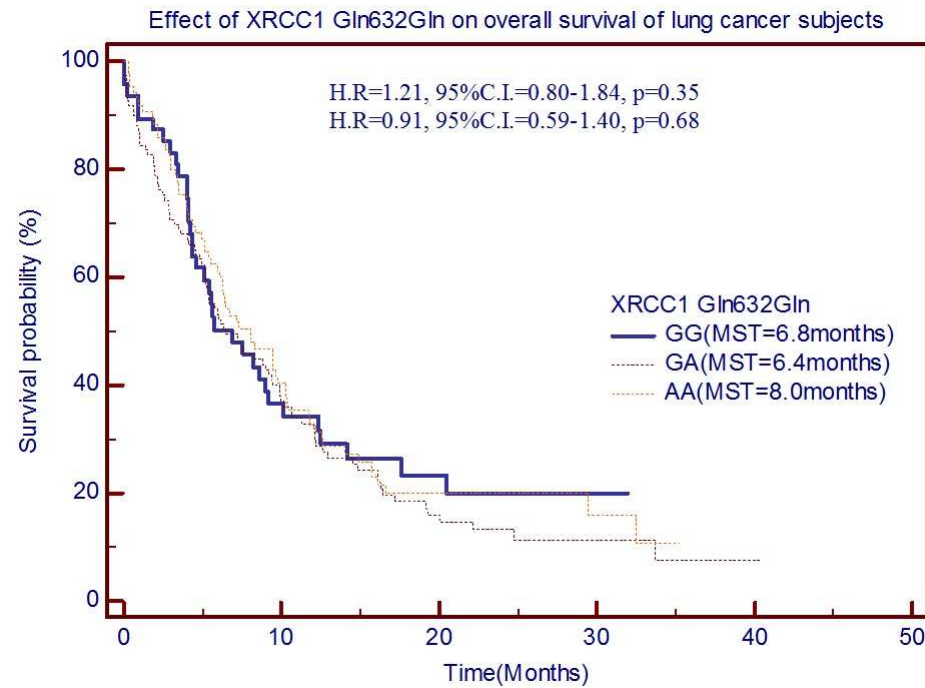
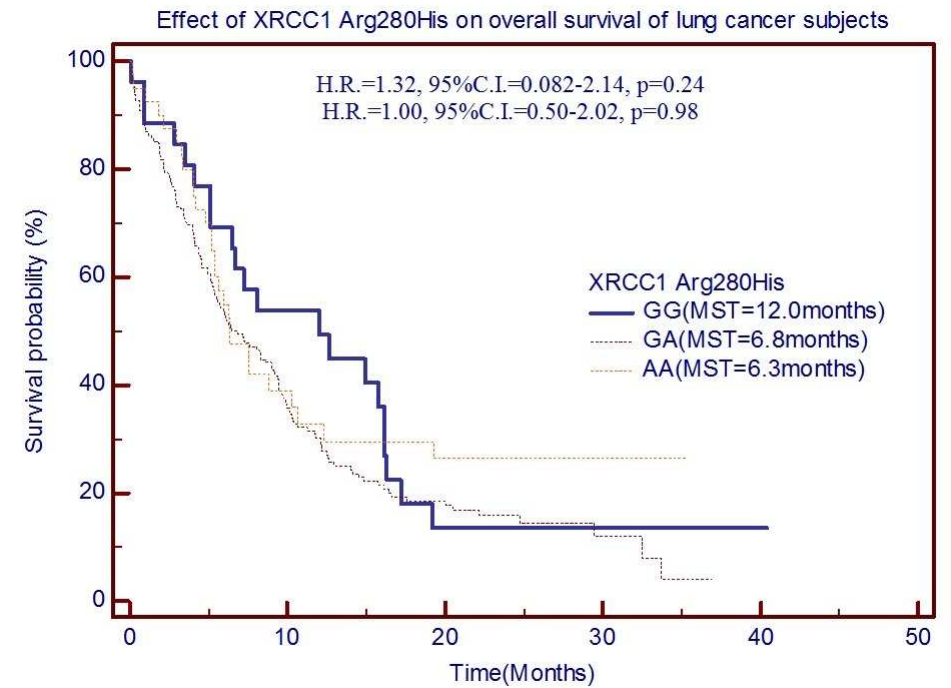
(a) *XRCC1 Gln⁶³²Gln*(b) *XRCC1 Gln²⁸⁰His*

Figure 5.21: Kaplan Meier Overall Survival Curve for (a) *XRCC1 Gln⁶³²Gln* (GG:reference) (b) *XRCC1 Gln²⁸⁰His* (GG:reference) for lung cancer patients. H.R:Hazard ration, MST:Median survival time

Table 5.47: *XRCC1* polymorphic variants and their association with overall survival

Genes	Cases N=241 n(%)	Death N=193 n(%)	Alive N=48 n(%)	Univariate analysis			Multivariate analysis	
				MST (months)	Log rank p-value	Unadjusted HR ^a	Adjusted HR ^b	p-value
XRCC1 Arg ³⁹⁹ Gln								
GG	68 (28.2)	58 (30.0)	10 (20.8)	6.3		1	1	
GA	130 (53.9)	102 (52.8)	28 (58.3)	6.4	0.93	0.98 (0.71-1.36)	1.03 (0.74-1.45)	0.82
AA	43 (17.8)	33 (17.0)	10 (20.8)	9.4	0.43	0.84 (0.56-1.29)	0.87 (0.56-1.35)	0.54
GA+AA	173 (71.7)	135 (69.9)	38 (79.1)	7.2	0.72	0.94 (0.69-1.29)	0.98 (0.72-1.34)	0.92
XRCC1 Arg ¹⁹⁴ Trp								
CC	190 (78.8)	154 (79.7)	36 (75)	7.2		1	1	
CT	50 (20.7)	39 (20.2)	11 (22.9)	5.9	0.97	0.99 (0.69-1.41)	0.9 (0.68-1.40)	0.90
TT	1 (0.41)	-	1 (2.0)	-		-	-	-
CT+TT	51 (21.2)	39 (20.2)	12 (25.0)	6.1	0.88	0.97 (0.68-1.37)	0.96 (0.67-1.38)	0.85
XRCC1 Pro ²⁰⁶ Pro								
AA	97 (40.2)	76 (39.3)	21 (43.7)	6.4		1	1	
AG	121 (50.2)	99 (51.2)	22 (45.8)	6.8	0.46	1.11 (0.82-1.50)	1.02 (0.75-1.39)	0.87
GG	23 (9.5)	18 (9.3)	5 (10.4)	8.8	0.72	1.09 (0.64-1.85)	1.15 (0.68-1.94)	0.60
AG+GG	144 (59.7)	117 (60.6)	27 (56.2)	7.3	0.45	1.114 (0.83-1.48)	1.03 (0.76-1.38)	0.83
XRCC1 Gln ⁶³² Gln								
GG	47 (19.5)	35 (18.1)	12 (25)	6.8		1	1	
GA	109 (45.2)	91 (47.1)	18 (37.5)	6.4	0.51	1.12 (0.7-1.66)	1.21 (0.80-1.84)	0.35
AA	85 (35.2)	67 (34.7)	18 (37.5)	8.0	0.90	0.98 (0.97-1.46)	0.91 (0.59-1.40)	0.68
GA+AA	194 (80.4)	158 (81.8)	36 (75)	7.2	0.74	1.06 (0.74-1.52)	1.06 (0.73-1.56)	0.73
XRCC1 Arg ²⁸⁰ His								
GG	26 (10.7)	21 (10.8)	5 (10.4)	12.0		1	1	
GA	175 (72.6)	144 (74.6)	31 (64.5)	6.8	0.22	1.31 (0.87-2.0)	1.32 (0.825-2.14)	0.24
AA	40 (16.5)	28 (14.5)	12 (25)	6.3	0.98	1.02 (0.57-1.77)	1.00 (0.50-2.02)	0.98
GA+AA	215 (89.2)	172 (89.1)	43 (89.5)	6.4	0.32	1.25 (0.82-1.90)	1.28 (0.80-2.05)	0.30

to be cont'd on next page

Table 5.47: *XRCC1* polymorphic variants and their association with overall survival (contd..)

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

5.10.3 Cox-regression analysis of the factors associated with overall survival of *XRCC1* polymorphic variants

The multivariate regression analysis showed that age sex, smoking, histology, stage, KPS were not associated with the prognosis for heterozygous subjects in case of polymorphic *XRCC1 Arg³⁹⁹Gln*. ECOG for subjects carrying heterozygous genotype in case of *XRCC1 Arg³⁹⁹Gln* was independent prognostic factor for the 3-year survival rate as depicted in table 5.48.

Table 5.48: Results of cox-regression multivariate analysis for subjects with polymorphic *XRCC1 Arg³⁹⁹Gln*

Variables	B	SE	Wald	P-value	Exp(b)	95%C.I.
XRCC1 Arg ³⁹⁹ Gln	0.037	0.172	0.047	0.827	1.038	0.742 - 1.451
AGE	0.0105	0.008	1.562	0.211	1.010	0.994- 1.027
SEX	0.032	0.299	0.012	0.912	1.033	0.576 - 1.853
Smoking	-0.047	0.279	0.028	0.865	0.953	0.552 - 1.645
Histology	0.073	0.120	0.377	0.539	1.076	0.851 -1.361
Stage	0.123	0.084	2.157	0.141	1.131	0.960 - 1.333
ECOG	0.401	0.163	6.013	0.014	1.493	1.085 - 2.054
KPS	0.089	0.186	0.231	0.630	1.093	0.760-1.572

B-Regression coefficient; SE-Standard error; Wald-Wald value;HR-Hazard ratio; 95% C.I.-confidence intervals

Table 5.49: Results of cox-regression multivariate analysis for subjects with polymorphic *XRCC1 Arg¹⁹⁴Trp*

Variables	B	SE	Wald	P-value	Exp(b)	95%C.I.
XRCC1 Arg ¹⁹⁴ Trp	-0.020	0.183	0.012	0.911	0.979	0.684 - 1.402
AGE	0.008	0.0078	1.142	0.285	1.008	0.993 - 1.024
SEX	-0.069	0.279	0.061	0.803	0.932	0.540 - 1.609
Smoking	-0.054	0.256	0.045	0.8312	0.946	0.574 - 1.560
Histology	0.045	0.106	0.186	0.665	1.047	0.850 - 1.288
Stage	0.0981	0.075	1.696	0.192	1.103	0.952 - 1.277
ECOG	0.405	0.148	7.418	0.006	1.499	1.122 - 2.004
KPS	0.038	0.169	0.052	0.818	1.039	0.747 - 1.446

B-Regression coefficient; SE-Standard error; Wald-Wald value;HR-Hazard ratio; 95% C.I.-confidence intervals

The multivariate regression analysis showed that age sex, smoking, histology, stage, KPS were not associated with the prognosis for heterozygous subjects in case of polymorphic *XRCC1 Arg¹⁹⁴ Trp*. ECOG for subjects carrying heterozygous genotype in case of *XRCC1 Arg¹⁹⁴ Trp* was independent prognostic factor for the 3-year survival rate as shown in table 5.49.

Table 5.50: Results of cox-regression multivariate analysis for subjects with polymorphic *XRCC1 Gln⁶³² Gln*

Variables	B	SE	Wald	P-value	Exp(b)	95%C.I.
XRCC1 Gln ⁶³² Gln	0.196	0.212	0.856	0.354	1.217	0.804 - 1.841
AGE	0.011	0.009	1.567	0.210	1.012	0.993 - 1.030
SEX	-0.463	0.363	1.623	0.202	0.629	0.309 - 1.278
Smoking	0.211	0.348	0.368	0.543	1.236	0.626 - 2.439
Histology	0.165	0.124	1.787	0.181	1.180	0.926 - 1.503
Stage	0.0728	0.095	0.583	0.444	1.075	0.893 - 1.295
ECOG	0.5109	0.180	8.055	0.004	1.666	1.173 - 2.367
KPS	-0.103	0.208	0.246	0.619	0.901	0.600 - 1.354

B-Regression coefficient; SE-Standard error; Wald-Wald value;HR-Hazard ratio; 95% C.I.- confidence intervals

The multivariate regression analysis showed that age sex, smoking, histology, stage, KPS were not associated with the prognosis for heterozygous subjects in case of polymorphic *XRCC1 Gln⁶³² Gln*. ECOG for subjects carrying heterozygous genotype in case of *XRCC1 Gln⁶³² Gln* was independent prognostic factor for the 3-year survival rate as shown in table 5.50.

5.10.4 Association of *XRCC1* polymorphisms and survival according to histological subtypes of lung cancer patients

The five polymorphic variants of *XRCC1* were analyzed for their prognostic relevance in predicting the prognosis of lung cancer patients with different histological subtypes. The data for each histological subtype has been demonstrated in different tables 5.51, 5.52 and 5.53 respectively. For the polymorphic variant *XRCC1 Arg³⁹⁹ Gln* subjects diagnosed with ADCC subtype, there were 21.6% subjects with mutant genotype (AA) and 49.3% subjects with heterozygous genotype (GA). Subjects with double allelic variant/mutant genotype (AA) showed an improved (MST= 9.4 months, log rank p-value=0.68) com-

pared to the subjects carrying heterozygous genotype (MST=7.6, log rank p-value=0.93) and wild type genotype (MST=6.3; figure 5.22a). Though there was no significant difference among the number of months survived, a trend was observed. Following univariate analysis, multivariate analysis consequently revealed subjects with single allelic variant (*GA*) genotype showed a visible trend towards high death rate compared to the wild type subjects (H.R.=1.7, 95%C.I.=0.873-3.35, p=0.12) whereas mutant genotype (*AA*) (H.R.=1.06, 95%C.I.=0.48-2.33, p=0.88) showed no modulation of death rate as shown in table 5.51. On the same lines subjects diagnosed with SCLC subtype and harboring mutant genotype (*AA*) showed a better MST (*AA*=10.4, log rank p-value=0.68) compared to other genotypes (MST: *GA*=6.7, *GG*=3.4, log rank p-value=0.2). Sequentially, the Cox proportional hazards ratio illustrated that the double allelic variants (*AA*) showed a trend of small death rate (H.R.=0.54, 95%C.I.=0.18-1.62, p=0.28) as shown in table 5.53.

For the polymorphic variant *XRCC1 Pro*²⁰⁶ the study observed that frequency distribution for subjects with ADCC subtype had 42.1% wild type subjects (*AA*), 50.6% heterozygous subjects (*AG*) and 7.2% mutant subjects (*GG*). There was no significant difference among the three genotypes of the SNPs. Study subjects diagnosed with SQCC subtype and carrying mutant genotype (*GG*) had a MST of 8.6 months (log rank p-value=0.98) whereas subjects with wild type and heterozygous genotype had a smaller MST (5.4 vs 7.5, log rank p-value=0.9). On the other hand no influence on death rate was observed for the SNP (*AG*: H.R.=0.99, 95%C.I.=1.61-1.60, p=0.98). SCLC subjects with heterozygous genotype (*AG*) showed a lower MST (MST=4.5, log rank p-value=0.39) compared to the mutant genotype (*GG*) (MST=6.5, log rank p-value=0.52), whereas wild type genotype (*AA*) showed a MST of 10.4 months as shown in figure 5.22b. The multivariate analysis suggested a trend of high death rate for the three genotypic combination (*GA*- H.R.=1.55, 95%C.I.=0.76-3.14, p=0.22: *AA*- H.R.=1.63, 95%C.I.=0.47-5.58, p=0.43: *GA+AA*- H.R.=1.45, 95%C.I.=0.74-2.82, p=0.27) as shown in table 5.52.

In case of *XRCC1 His*⁶³²*His* for patients with ADCC subtype and carrying single allelic variant genotype (*GA*) showed the highest MST of 9.2 months (Log rank p-value=0.55) as compared to subjects with wild type (*GG*) (MST=5.7) and mutant genotype (*AA*) (MST=6.2, log rank p value=0.50). Subjects with SCLC subtype and carrying heterozygous genotype (*GA*) showed a MST of 5.8 months (Log rank p-value=0.33) with a high death rate (H.R.=3.08, 95%C.I.=1.17-8.10, **p=0.02**). The double allelic variant carriers (*AA*) showed a MST of 8.5 months (Log rank p-value=0.72) (Figure 5.22c) with no significant relation with death rate as shown in table 5.53.

The next SNP following the sequence is *XRCC1 280* where subjects with wild type

genotype (*GG*) and diagnosed with ADCC showcased the highest MST of 14.9 months whereas single allelic variants (*GA*) (MST=7.6 months, log rank p-value=0.22) and double allelic variant genotype (*AA*) (MST=7.5, log rank p-value=0.70) showed similar time of survival. On adjusting for factors sex, age, smoking, stage, KPS, ECOG, heterozygous subjects revealed a trend of high death rate for lung cancer subjects (H.R.=1.81, 95%C.I.=0.68-4.85, p=0.23). Subjects with SQCC and SCLC subtype and carrying polymorphic variant *XRCC1 280* showed no effect on survival as shown in table 5.52 and table 5.53

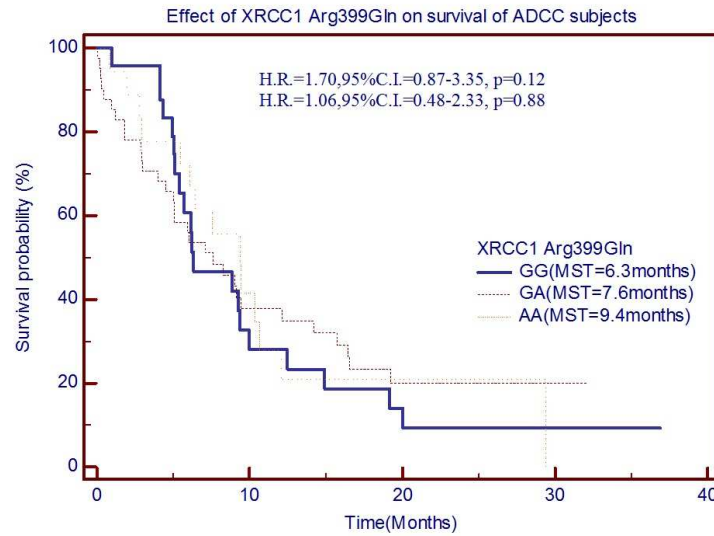
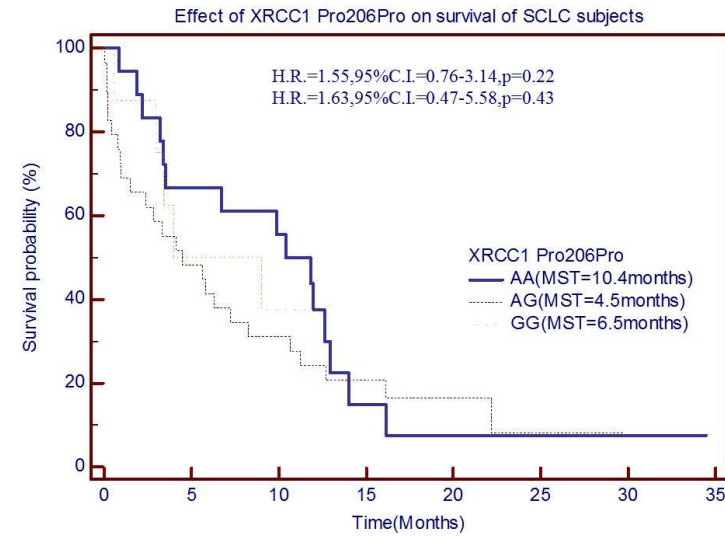
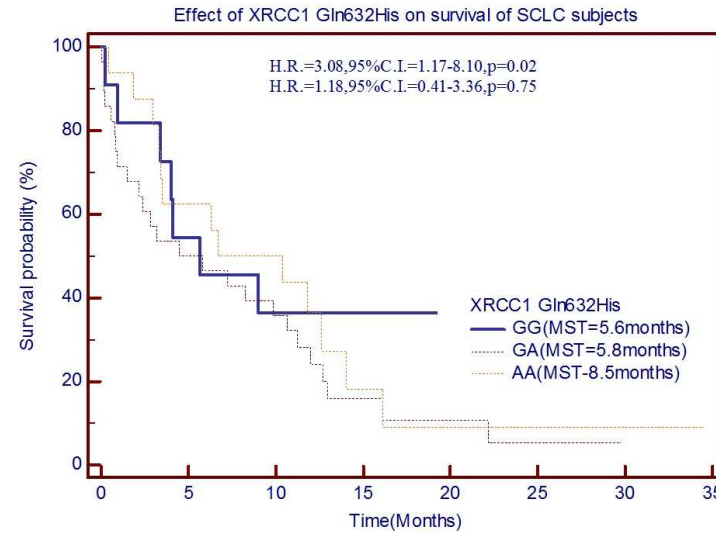
(a) *XRCC1 Arg³⁹⁹Gln*(b) *XRCC1 Pro²⁰⁶Pro*(c) *XRCC1 Gln⁶³²Gln*

Figure 5.22: Kaplan Meier survival Curves for *XRCC1* variants in lung cancer patients stratified according to histology (a) *Arg³⁹⁹Gln* (GG:reference) in ADCC patients (b) *Pro²⁰⁶Pro* (AA:reference) in SCLC patients (c) *Gln⁶³²Gln* (GG:reference) in SCLC patients. H.R:Hazard ration, MST:Median survival time

Table 5.51: Association of *XRCC1* genetic variants with survival of lung cancer patients according to Tumor histology ADCC

Genes	ADCC					
XRCC1 ³⁹⁹ G/A (N)	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	24 (8.9)	6.3		1	1	
GA	41 (49.3)	7.6	0.93	0.90 (0.51-1.6)	1.70 (0.87-3.35)	0.12
AA	18 (21.6)	9.4	0.68	0.89 (0.44-1.70)	1.06 (0.48-2.33)	0.88
GA+AA	59 (71.08)	8.3	0.70	0.90 (0.52-1.54)	0.70 (0.33-1.47)	0.35
XRCC1 ¹⁹⁴ C/T (N)	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	63 (75.9)	8.8		1	1	
CT	19 (22.8)	6.1	0.91	1.03 (0.56-1.87)	0.89 (0.46-1.72)	0.74
TT	1 (1.2)	-		-	-	-
CT+TT	20	6.1	0.94	0.97 (0.54-1.7)	0.87 (0.45-1.67)	0.68
XRCC1 ²⁰⁶ A/G (N)	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	35 (42.1)	7.6		1	1	
AG	42 (50.6)	7.5	0.62	1.13 (0.68-1.87)	0.85 (0.48-1.49)	0.57
GG	6 (7.2)	8.8	0.84	1.09 (0.40-2.9)	1.07 (0.35-3.19)	0.90
AG+GG	48 (57.8)	8.8	0.60	1.13 (0.69-1.85)	0.88 (0.51-1.51)	0.65
XRCC1 ⁶³² G/A (N)	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	18 (21.6)	5.7		1	1	
GA	38 (45.7)	9.2	0.55	0.82 (0.41-1.62)	0.61 (0.26-1.4)	0.25
AA	27 (32.5)	6.2	0.50	1.25 (0.65-2.41)	1.26 (0.52-3.00)	0.60
GA+AA	65 (78.3)	8.8	0.97	0.99 (0.53-1.82)	0.88 (0.43-1.78)	0.72
XRCC1 ²⁸⁰ G/A (N)	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	8 (9.6)	14.9		1	1	
GA	60 (72.2)	7.6	0.22	1.74 (0.83-3.6)	1.81 (0.68-4.85)	0.23
AA	15 (18.0)	7.5	0.70	1.23 (0.43-3.48)	0.78 (0.19-3.20)	0.73
GA+AA	75 (90.3)	7.56	0.28	1.63 (0.77-3.4)	1.72 (0.64-4.58)	0.27

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.52: Association of *XRCC1* genetic variants with survival of lung cancer patients according to Tumor histology SQCC

Genes	SQCC					
XRCC1 ³⁹⁹ G/A (N)	Cases n(%) N=101	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	30 (29.7)	10.1		1	1	

to be cont'd on next page

Table 5.52: Association of *XRCC1* genetic variants with survival of lung cancer patients according to Tumor histology SQCC (cont..)

Genes	SQCC					
GA	57 (56.4)	5.4	0.41	1.2 (0.76-1.93)	1.26 (0.73-2.20)	0.39
AA	14 (13.8)	9.4	0.37	0.70 (0.35-1.44)	0.65 (0.29-1.45)	0.30
GA+AA	71 (70.2)	6.4	0.74	1.70 (0.68-1.70)	0.95 (0.58-1.55)	0.84
XRCC1 ¹⁹⁴ C/T (N)	Cases n(%) N=101	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	82 (81.1)	6.4		1	1	
CT	19 (18.8)	8.0	0.86	0.95 (0.55-1.62)	0.89 (0.50-1.58)	0.69
TT		-		-	-	-
CT+TT				-	-	-
XRCC1 ²⁰⁶ A/G (N)	Cases n(%) N=101	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	44 (43.5)	5.4		1	1	
AG	48 (47.5)	7.5	0.9	0.97 (0.61-1.53)	0.99 (1.61-1.60)	0.98
GG	9 (8.9)	8.6	0.98	1.01 (0.47-2.19)	-	-
AG+GG	57 (56.4)	8.0	0.94	0.98 (0.63-1.53)	1.00 (0.63-1.59)	0.97
XRCC1 ⁶³² G/A (N)	Cases n(%) N=101	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	18 (17.8)	6.8		1	1	
GA	43 (42.5)	5.2	0.5	1.20 (0.69-2.18)	1.17 (0.59-2.73)	0.63
AA	40 (39.6)	8.0	0.39	0.77 (0.39-1.47)	0.65 (0.32-1.31)	0.23
GA+AA	83 (82.1)	6.4	0.92	0.97 (0.55-1.71)	0.88 (0.48-1.61)	0.68
XRCC1 ²⁸⁰ G/A (N)	Cases n(%) N=101	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	10(9.9)	15.7		1	1	
GA	73(72.2)	6.4	0.64	1.15 (0.61-2.28)	1.34 (0.62-2.88)	0.44
AA	18 (17.8)	5.3	0.84	0.90 (0.39-2.16)	0.77 (0.27-2.19)	0.63
GA+AA	91 (90.0)	6.4	0.78	1.10 (0.56-2.15)	1.25 (0.59-2.61)	0.54

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.53: Association of *XRCC1* genetic variants with survival of lung cancer patients according to Tumor histology SCLC

Genes	SCLC					
XRCC1 ³⁹⁹ G/A (N)	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	14 (25.4)	3.4		1	1	
GA	32 (58.1)	6.7	0.2	0.66 (0.31-1.4)	1.00 (0.40-2.46)	0.99
AA	9 (16.3)	10.4	0.68	0.84 (0.35-1.9)	0.54 (0.18-1.62)	0.28
GA+AA	41 (74.5)	7.2	0.25	0.68 (0.32-1.41)	0.70 (0.33-1.47)	0.35

to be cont'd on next page

Table 5.53: Association of *XRCC1* genetic variants with survival of lung cancer patients according to Tumor histology SCLC (cont..)

Genes	SCLC					
	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
XRCC1¹⁹⁴C/T (N)		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	44 (80.0)	7.2		1	1	
CT	11 (20.2)	3.5	0.68	1.16 (0.54-2.5)	0.97 (0.41-2.30)	0.95
TT	-	-	-	-	-	
CT+TT	-	-	-	-	-	
XRCC1²⁰⁶A/G (N)	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	18 (32.7)	10.4		1	1	
AG	29 (52.7)	4.5	0.39	1.32 (0.171-2.47)	1.55 (0.76-3.14)	0.22
GG	8 (14.5)	6.5	0.52	1.08 (0.45-4.12)	1.63 (0.47-5.58)	0.43
AG+GG	37 (67.2)	4.5	0.43	1.27 (0.70-2.33)	1.45 (0.74-2.82)	0.27
XRCC1⁶³²G/A (N)	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	11 (20.0)	5.6		1	1	
GA	28 (50.9)	5.8	0.33	1.54 (0.70-3.21)	3.08 (1.17-8.10)	0.02
AA	16 (29.0)	8.5	0.72	1.15 (0.48-2.90)	1.18 (0.41-3.36)	0.75
GA+AA	44 (80)	6.7	0.42	1.38 (0.67-2.86)	1.92 (0.81-4.56)	0.13
XRCC1²⁸⁰G/A (N)	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	8 (14.5)	9.3		1	1	
GA	41(74.5)	5.8	0.64	1.20 (0.56-2.6)	1.00 (0.39-2.52)	0.99
AA	6 (10.9)	5.9	0.31	1.54 (0.49-5.9)	0.97 (0.13-6.86)	0.98
GA+AA	75 (90.3)	47 (85.4)	0.58	1.24 (0.59-2.63)	1.02 (0.41-2.53)	0.95

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

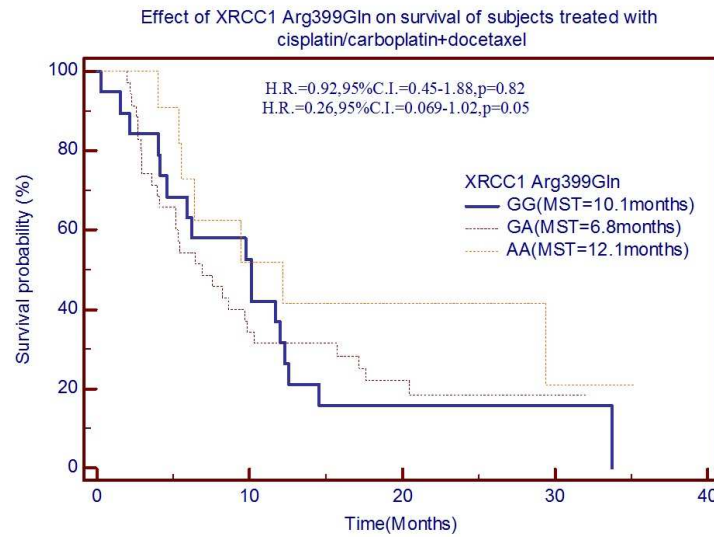
5.10.5 Association of genetic determinants of *XRCC1* and overall survival of lung cancer patients treated with different chemotherapeutic regimens

Tables 5.54, 5.55, 5.56 below shows the effect of the five SNPs of *XRCC1* on survival after stratifying for the chemotherapeutic regimens used for the treatment of lung cancer patients. Chemotherapy was given to 165 patients and three regimens were administered to different sets of patients. The three regimens used included a platinum based drug

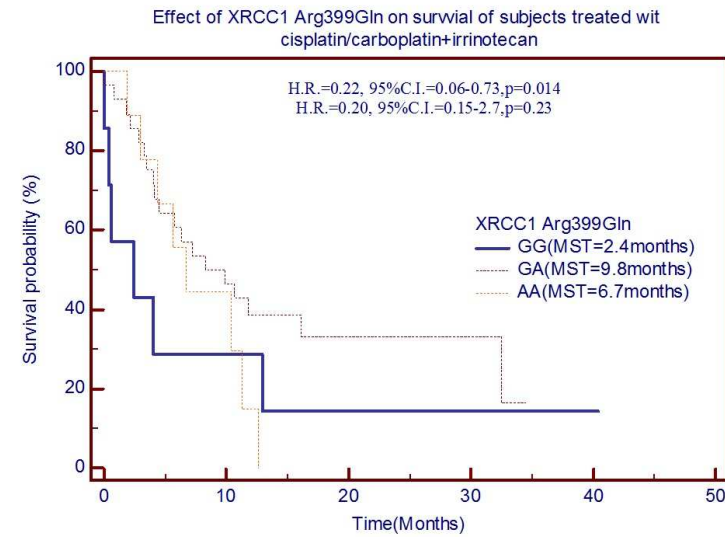
(cisplatin and carboplatin) which was common to all lung cancer patients, along with this patients also received a second line chemotherapeutic agent like docetaxel, irinotecan, or pemetrexed. MST (median survival time) was evaluated for the five polymorphic variant using univariate analysis while multivariate analysis evaluated the hazards rate respectively.

As shown in table 5.54 it was observed that patients who were receiving cisplatin/ carboplatin along with docetaxel and carrying the mutant genotype (*AA*) of *XRCC1 Arg³⁹⁹ Gln* gene had a longer survival time with MST of 12.1 months figure (5.23a) and a significant small death rate (H.R.=0.26, 95%C.I.=0.069-1.02, ***p*=0.05**) indicating a better survival. Single allelic carriers (*GA*) for the same gene showed a lower survival time of 6.8 months and hazards rate of 1 fold (H.R.=0.92, 95%C.I.=0.45-1.88). Further, when lung cancer patients with the heterozygous genotype (*GA*) of *XRCC1 Arg³⁹⁹ Gln* gene were administered irinotecan+cisplatin/carboplatin, the MST was around 9.8 months which was the highest among both mutant (*AA*) (MST=6.7 months, Log rank p-value=0.87) as shown in table 5.55 and combined variant (*GA+AA*)(MST=8.2 months, log rank p-value=0.23) genotype (figure 5.23b) Multivariate analysis after applying cox regression model revealed a significant better survival for subjects with single allelic variant genotype (*GA*) (H.R=0.22, 95%C.I.=0.06-0.73, ***p*=0.014**) as compared to the subjects being administered same drug regime but with mutant genotype (*AA*) (MST=6.7, HR=0.20, 95%C.I.=0.15-2.70, p=0.23).

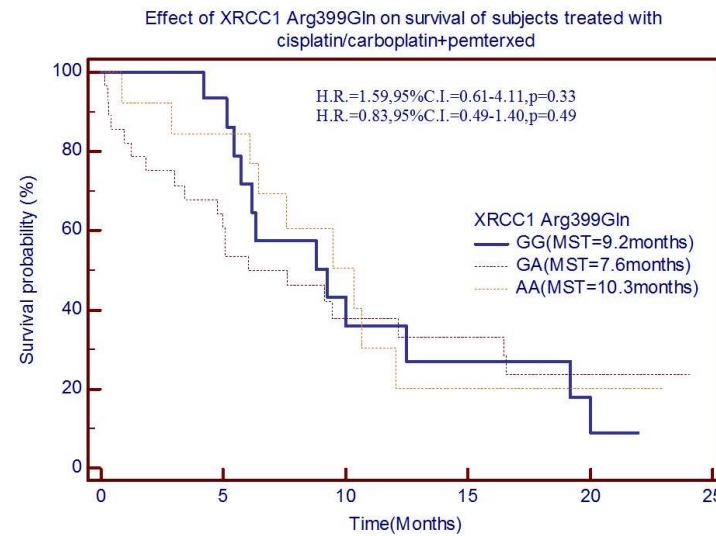
Study subjects harbouring the mutant genotype (*AA*) in case of *XRCC1 Arg³⁹⁹ Gln* and given pemetrexed+cisplatin/carboplatin, the MST was around 10.3 months (figure 5.23c) with no association with death rate (H.R.=0.83, 95%C.I.=0.49-1.40, p=0.49) as shown in table 5.56. Whereas, subjects with heterozygous genotype (*GA*) showed a trend of smaller (MST =7.6 months, Log rank p-value=0.81) and sequentially a high death rate (H.R.=1.59, 95%C.I.=0.61-4.11, p=0.33).



(a) Cisplatin/Carboplatin+Docetaxel

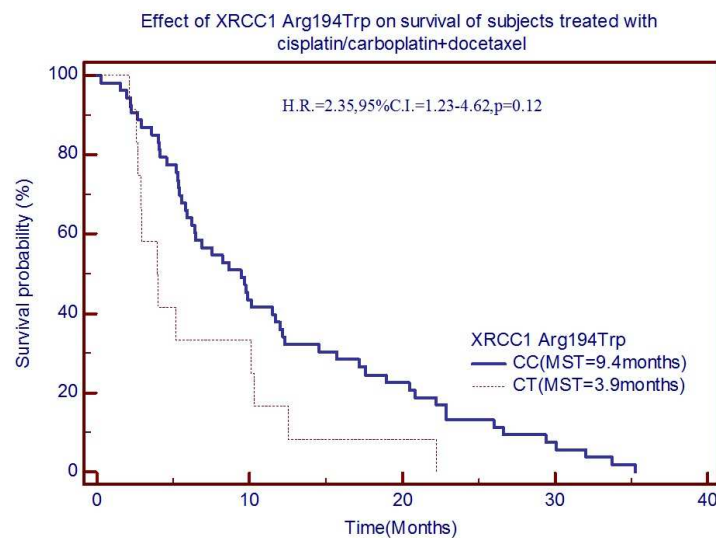


(b) Cisplatin/Carboplatin+Irinotecan

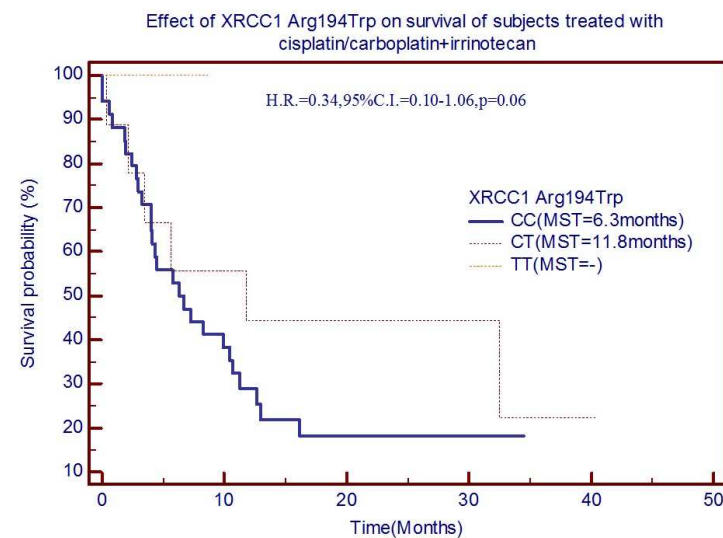


(c) Cisplatin/Carboplatin+Pemetrexed

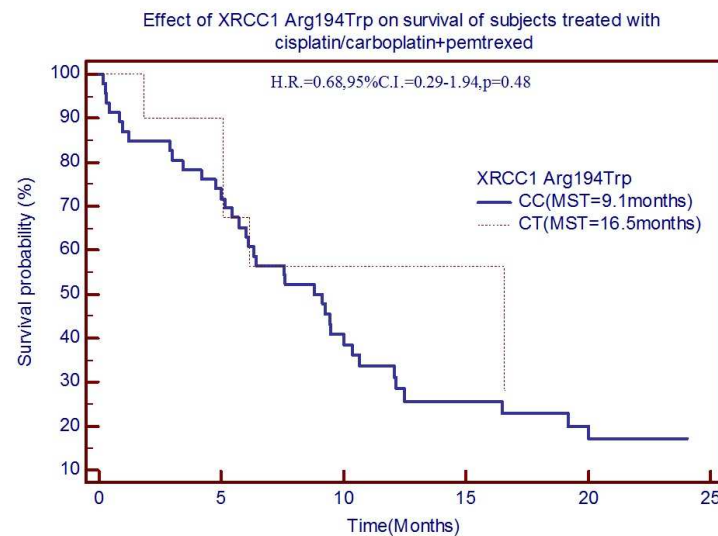
Figure 5.23: Kaplan Meier Overall Survival Curve for *XRCC1* Arg³⁹⁹Gln treated with (a) Cisplatin/Carboplatin+Docetaxel (b) Cisplatin/Carboplatin+Irinotecan (c) Cisplatin/Carboplatin+Pemetrexed



(a) Cisplatin/Carboplatin+Docetaxel



(b) Cisplatin/Carboplatin+Irinotecan



(c) Cisplatin/Carboplatin+Pemetrexed

Figure 5.24: Kaplan Meier Overall Survival Curve for *XRCC1* Arg¹⁹⁴Trp treated with (a) Cisplatin/Carboplatin+Docetaxel (b) Cisplatin/Carboplatin+Irinotecan (c) Cisplatin/Carboplatin+Pemetrexed

In case of *XRCC1 Arg¹⁹⁴ Trp*, subjects carrying at least one dominant T allele (*CT*) and receiving cisplatin/carboplatin along with docetaxel depicted a shorter MST of 3.9 months (Log rank test $p=0.02$), figure 5.24a and a significant higher death rate (H.R.=2.35, 95%C.I.=1.23-4.62, $p=0.012$). On the contrary patients who were heterozygotes (*CT*) for the *Arg¹⁹⁴ Trp* polymorphic site and receiving irinotecan+cisplatin/carboplatin showed a higher survival with MST of 11.8 months (log-rank test $p=0.35$, figure 5.24b) as shown in table 5.55. In the multivariate cox proportional hazards model, after adjusting for various factors a nearly significant shorter death rate was seen in the same subjects (H.R.=0.34, 95%C.I.=0.10-1.06, $p=0.06$).

Subjects for the next snp *XRCC1 Pro²⁰⁶ Pro* and administered cisplatin/carboplatin along with docetaxel did not show any significant difference between the median survival time for the three genotypic combinations (*AA*-6.4, *AG*-9.8, log rank p -value=0.85; *GG*-8.6, log rank p -value=0.71) as indicated in table 5.55. Hazard rate did not show any significant change in relation to lung cancer. Similarly, subjects administered irinotecan and pemetrexed did not show any significant change in survival time or death rate.

For the polymorphic variant *XRCC1 His⁶³²*, subjects given cisplatin/carboplatin along with docetaxel and carrying heterozygous genotype (*GA*) showed a MST of 7.5 months (Log rank p -value=0.32) with a trend of high death rate (H.R.=1.69, 95%C.I.=0.68-4.18, $p=0.25$). Subjects carrying mutant genotype (*AA*) under the same drug administration showed a better MST (9.7 months, log rank p -value=0.82) though there was no significant difference between MST for the three genotypic combinations as shown in table 5.54.

Subjects treated with cisplatin/carboplatin along with pemetrexed and carrying single allele of variant genotype (*GA*) for *XRCC1 His⁶³²* showed a better MST (10.6, log rank p -value=0.17) compared to the wild (*GG*) (MST=5.7) and mutant (*AA*) (MST=6.1, log rank p -value=0.88) genotype as shown in table 5.56. After cox-regression analysis, the death rate was smaller for single allelic carriers (*GA*) (H.R.=0.47, 95%C.I.=0.16-1.37, $p=0.17$) compared to the double allelic variants (*AA*) (H.R.=1.54, 95%C.I.=0.48-4.59, $p=0.46$).

The next important polymorphic variant of *XRCCC1* studied was *XRCC1 Arg²⁸⁰*. Our data showed that subjects with the wild type (*GG*) genotype of the polymorphism and undergoing the chemotherapy with cisplatin/ carboplatin and docetaxel showed the highest survival time of 13.8 months compared to MST of 8.2 (Log rank p -value=0.9) and 7.5 months (Log rank p -value=0.69) for heterozygous (*GA*) and mutant (*AA*) genotypic subjects respectively. The Cox-regression analysis did not show any association with hazards rate.

Table 5.54: Association of *XRCC1* genetic variants with chemotherapy regimen (Cisplatin/Carboplatin+Docetaxel) and survival

Genes	Cisplatin/Carboplatin+Docetaxel					
XRCC1 ³⁹⁹ G/A	Cases n(%) N=65	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	19 (29.2)	10.1		1	1	
GA	35 (53.8)	6.8	0.92	1.03 (0.55-1.93)	0.92 (0.45-1.88)	0.82
AA	11 (16.9)	12.1	0.2	0.58 (0.27-1.27)	0.26 (0.069-1.02)	0.05
GA+AA	46 (70.7)	7.5	0.71	0.89 (0.49-1.62)	0.76 (0.38-1.51)	0.44
XRCC1 ¹⁹⁴ C/T	Cases n(%) N=65	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	53 (81.5)	9.4		1	1	
CT	12 (18.4)	3.9	0.02	1.98 (0.88-1.44)	2.35 (1.23-4.62)	0.12
TT	-	-		-	-	
CT+TT	-	-		-	-	
XRCC1 ²⁰⁶ A/G	Cases n(%) N=65	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	31 (47.6)	6.4		1	1	
AG	27 (41.5)	9.8	0.85	0.94 (0.53-1.69)	0.97 (0.49-1.92)	0.95
GG	7 (10.7)	8.6	0.71	1.18 (0.46-3.02)	1.01 (0.34-3.0)	0.98
AG+GG	34 (52.3)	9.4	0.97	0.99 (0.57-1.70)	1.03 (0.54-1.97)	0.90
XRCC1 ⁶³² G/A	Cases n(%) N=65	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	15 (23.0)	8.6		1	1	
GA	23 (35.3)	7.5	0.32	1.42 (0.70-2.8)	1.69 (0.68-4.18)	0.25
AA	27(41.5)	9.7	0.82	0.92 (0.44-1.90)	1.07 (0.46-2.47)	0.86
GA+AA	50 (76.9)	9.4	0.76	1.10 (0.58-2.07)	1.17 (0.57-2.37)	0.66
XRCC1 ²⁸⁰ G/A	Cases n(%) N=65	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	4 (6.1)	13.8		1	1	
GA	49 (75.3)	8.2	0.9	1.06(0.39-2.9)	1.01 (0.29-3.4)	0.97
AA	12(18.4)	7.5	0.69	0.78(0.22-2.7)	-	-
GA+AA	61 (93.8)	8.2	0.98	1.00(0.36-2.7)	1.01 (0.30-3.34)	0.98

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.55: Association of *XRCC1* genetic variants with chemotherapy regimen (Cisplatin/ Carboplatin+Irinotecan) and survival

Genes	Cisplatin/Carboplatin+Irinotecan					
XRCC1 ³⁹⁹ G/A	Cases n(%) N=44	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	7 (15.9)	2.4		1	1	

to be cont'd on next page

Table 5.55: Association of *XRCC1* genetic variants with chemotherapy regimen (Cisplatin/ Carboplatin+Irinotecan) and survival (cont..)

Genes	Cisplatin/Carboplatin+Irinotecan					
GA	28 (63.6)	9.8	0.15	0.52 (0.17-1.6)	0.22 (0.06-0.73)	0.014
AA	9 (20.4)	6.7	0.87	0.92 (0.31-2.68)	0.20 (0.15-2.7)	0.23
GA+AA	37 (84.0)	8.2	0.23	0.59 (0.20-1.7)	0.23 (0.078-0.72)	0.012
XRCC1¹⁹⁴C/T	Cases n(%) N=44	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	34 (77.2)	6.3		1	1	
CT	9 (20.4)	11.8	0.35	0.66 (0.30-1.46)	0.34 (0.10-1.06)	0.06
TT	1 (2.2)	-		-	-	
CT+TT		11.8	0.25	0.60 (0.28-1.30)	0.32 (0.10-1.01)	0.05
XRCC1²⁰⁶A/G	Cases n(%) N=44	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	14 (31.8)	10.4		1	1	
AG	24 (54.5)	6.3	0.57	1.22 (0.59-2.53)	1.42 (0.60-3.36)	0.46
GG	6 (13.6)	-	0.97	1.01 (0.28-3.67)	1.34 (0.29-6.13)	0.70
AG+GG	30 (68.1)	6.3	0.65	1.17 (0.58-2.38)	1.22 (0.54-2.77)	0.62
XRCC1⁶³²G/A	Cases n(%) N=44	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	8 (18.1)	4.8		1	1	
GA	20 (45.4)	8.2	0.75	1.17 (0.44-3.07)	5.57 (1.17-26.5)	0.03
AA	16 (36.3)	6.5	0.95	0.96 (0.33-2.77)	1.13 (0.28-4.59)	0.85
GA+AA	36 (81.8)	7.2	0.85	1.09 (0.43-2.7)	1.75 (0.62-4.93)	0.28
XRCC1²⁸⁰G/A	Cases n(%) N=44	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	8 (18.1)	12.6		1	1	
GA	30 (68.1)	7.23	0.24	1.73 (0.75-3.98)	1.09 (0.35-3.34)	0.87
AA	6 (13.6)	5.9	0.12	2.43 (0.62-9.53)	0.93 (0.09-8.9)	0.95
GA+AA	36 (81.8)	5.8	0.19	1.82 (0.82-4.02)	1.26 (0.43-3.70)	0.66

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.56: Association of *XRCC1* genetic variants with chemotherapy regimen (Cisplatin/ Carboplatin+Pemetrexed) and survival

Genes	Cisplatin/Carboplatin+Pemetrexed					
XRCC1³⁹⁹G/A	Cases n(%) N=56	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	15 (26.7)	9.2		1	1	
GA	28 (50)	7.6	0.81	1.07 (0.53-2.2)	1.59 (0.61-4.11)	0.33
AA	13 (23.2)	10.3	0.86	0.92 (0.39-2.18)	0.83 (0.49-1.40)	0.49
GA+AA	41 (73.2)	9.2	0.94	1.02	1.19 (0.55-2.5)	0.65

to be cont'd on next page

Table 5.56: Association of *XRCC1* genetic variants with chemotherapy regimen (Cisplatin/ Carboplatin+Pemetrexed) and survival ((cont..)

Genes	Cisplatin/Carboplatin+Pemetrexed					
XRCC1 ¹⁹⁴ C/T	Cases n(%) N=56	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	46 (82.1)	9.1		1	1	
CT	10 (17.8)	16.5	0.60	0.78 (0.33-1.84)	0.68 (0.29-1.94)	0.48
TT	-	-		-	-	
CT+TT	-	-		-	-	
XRCC1 ²⁰⁶ A/G	Cases n(%) N=56	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	23 (41.0)	9.4		1	1	
AG	27 (48.2)	9.1	0.76	1.10 (0.57-2.14)	1.00 (0.44-2.33)	0.99
GG	6 (10.7)	7.2	0.26	1.67(0.56-4.9)	1.34 (0.75-2.38)	0.31
AG+GG	33 (58.9)	8.8	0.55	1.20 (0.65-2.24)	1.22 (0.59-2.5)	0.58
XRCC1 ⁶³² G/A	Cases n(%) N=56	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	10 (17.8)	5.7		1	1	1
GA	29 (51.7)	10.6	0.17	0.57 (0.22-1.48)	0.47 (0.16-1.37)	0.17
AA	17 (30.3)	6.1	0.88	1.06 (0.44-2.55)	1.54 (0.48-4.95)	0.46
GA+AA	46 (82.1)	9.4	0.39	0.71 (0.30-1.70)	0.90 (0.36-2.24)	0.82
XRCC1 ²⁸⁰ G/A	Cases n(%) N=56	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	6 (10.7)	12.8		1	1	
GA	37 (66.0)	9.4	0.46	1.43 (0.57-3.58)	1.48 (0.50-4.3)	0.47
AA	13 (23.2)	7.5	0.73	1.22 (0.38-3.9)	0.56 (0.1-2.70)	0.47
GA+AA	50 (89.2)	9.1	0.5	1.38 (0.55-3.43)	1.23 (0.42-3.54)	0.70

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.10.6 Genetic determinants of *XRCC1* and their correlation with clinical stage and performance status in relation to survival

To elucidate further if performance status and stage behave as independent prognostic factors the two performance factors ECOG, KPS and stage were investigated for the five SNPs, for their association with overall survival. Median survival time and hazard rate was evaluated for every genotypic combination for the five polymorphic sites of

XRCC1.

Upon investigation our results suggested that none of the genotypic combination for the five SNPs suggested modulation of the survival of lung cancer patients with relevance to stage as shown in table 5.57 and 5.58. For *XRCC1 Arg³⁹⁹ Gln* in case of stage III patients, there were 13.8% subjects with mutant genotype (*AA*), 57.6% heterozygous subjects and 28.4% subjects with wild genotype (*GG*). A higher MST of 11.2 months was observed for mutant genotype (*AA*) in comparison to heterozygous (*GA*) (MST=6.8) and wild type (*GG*) (MST=9.3) genotype (Log rank p-value=0.92). Multivariate analysis for the polymorphic site showed no significant relation with death rate. In case of *XRCC1 Arg³⁹⁹ Gln* subjects diagnosed with stage IV as shown in table 5.58 exhibited there were 49.5% single allelic variants (*GA*), 22.5% double allelic variants (*AA*) and 72.0% of subjects carried combined variant genotype (*GA+AA*) compared to 27.9% for wild type genotype (*GG*). MST of 6.3 months was observed for single allelic carriers (*GA*), 6.7 months for mutant subjects (*AA*) and MST of 6.4 months was observed for combined variants (*GA+AA*) respectively. Cox-regression analysis for heterozygous (*GA*) subjects indicated a trend towards increased death rate (H.R.=1.38, 95%C.I.=0.81-2.34, p=0.23). Whereas subjects with combined variant genotype (*GA+AA*) revealed a hazard rate of 1.25 fold (H.R.=1.25, 95% C.I.=0.77-2.05, p=0.36). The polymorphic variants *XRCC1 Arg¹⁹⁴ Trp*, *Pro²⁰⁶ Pro*, *Gln⁶³² Gln* and *Arg²⁸⁰ His* did not modulate lung cancer survival with respect to stage IV of subjects as shown in table 5.58.

For subjects with stage III tumor the genetic variants of *XRCC1 Arg³⁹⁹ Gln*, *Arg¹⁹⁴ Trp*, *Pro²⁰⁶ Pro*, and *Gln⁶³² Gln* did not show any association with lung cancer survival. In case of *XRCC1 Arg²⁸⁰ Gln* it was observed for patients with stage III tumor, 10.7% subjects harboured wild type (*GG*) genotype. The frequency distribution further suggested there were 70.0% and 19.2% heterozygous and mutant subjects respectively. Survival analysis suggested a MST of 14.9 months for subjects (*GG*) with wild type genotype which was the highest compared to median survival time of subjects with heterozygous (*GA*) (MST=8.2) and mutant (*AA*) (MST=6.3) (Log rank p-value=0.54) genotype though the difference between them is not statistically significant. For subjects with combined variant (*GA+AA*) genotype a trend towards high death rate was observed (H.R.=1.62, 95%C.I.=0.79-3.34, p=0.18) after cox-regression analysis.

Table 5.57: Association of Polymorphic *XRCC1* variants with survival of lung cancer patients according to pathological stage IV

Genes	Stage IV					
XRCC1 ³⁹⁹ G/A	Cases n(%) N=111	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
	GG	31 (27.9)	6.2	0.66	1	
	GA	55 (49.5)	6.3		1.01	1.38 (0.81-2.34)
	AA	25 (22.5)	6.7		0.80	0.85 (0.43-1.67)
GA+AA	80 (72.0)	6.4		0.93	1.25 (0.77-2.05)	0.36
XRCC1 ¹⁹⁴ C/T	Cases n(%) N=111	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
	CC	90 (81.8)	7.1	0.53	1	
	CT	21 (18.9)	5.4		1.17	1.11 (0.65-1.80)
	TT	0	-		-	-
CT+TT	21 (18.9)	6.4		0.93	1.25(0.77-2.05)	0.36
XRCC1 ²⁰⁶ A/G	Cases n(%) N=111	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
	AA	44 (39.6)	6.1	0.9	1	
	AG	57 (51.3)	6.4		1.10	0.94 (0.58-1.50)
	GG	10 (9.0)	9.0		1.11	1.01 (0.46-2.20)
AG+GG	67 (60.3)	7.3		1.10	0.93 (0.59-1.45)	0.76
XRCC1 ⁶³² G/A	Cases n(%) N=111	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
	GG	21 (18.9)	5.6	0.93	1	
	GA	51 (45.9)	6.4		1.10	1.11 (0.60-2.05)
	AA	39 (35.1)	6.46		1.05	1.19 (0.60-2.35)
GA+AA	90 (81.0)	6.4		1.08	1.16 (0.65-2.06)	0.60
XRCC1 ²⁸⁰ G/A	Cases n(%) N=111	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
	GG	12 (10.8)	7.2	0.49	1	
	GA	84 (75.6)	6.1		1.19	1.15 (0.60-2.02)
	AA	15 (13.5)	6.3		0.82	0.81 (0.30-2.16)
GA+AA	99 (89.1)	6.1		1.13	1.05 (0.55-1.99)	0.88

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.58: Association of Polymorphic *XRCC1* variants with survival of lung cancer patients according to pathological stage III

Genes	Stage III					
XRCC1 ³⁹⁹ G/A	Cases n(%) N=130	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	37 (28.4)	9.3	0.92	1	1	
GA	75 (57.6)	6.8		0.98	0.83 (0.52-1.33)	0.45
AA	18 (13.8)	11.2		0.88	0.67 (0.32-1.37)	0.28
GA+AA	93 (71.5)	7.6		0.96	0.85 (0.54-1.34)	0.49
XRCC1 ¹⁹⁴ C/T	Cases n(%) N=130	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
CC	100 (76.9)	7.5	0.62	1	1	
CT	29 (22.3)	10.2		0.89	0.90 (0.54-1.48)	0.69
TT	1 (0.7)	-		-	-	-
CT+TT	30 (23.0)	-	-	-	0.91 (0.54-1.49)	0.71
XRCC1 ²⁰⁶ A/G	Cases n(%) N=130	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
AA	53 (40.7)	9.6	0.84	1	1	
AG	64 (49.2)	7.2		1.12	1.13 (0.74-1.74)	0.54
GG	13 (23.0)	8.6		1.07	1.39 (0.65-2.98)	0.39
AG+GG	77 (59.2)	7.5		1.11	1.13 (0.75-1.70)	0.54
XRCC1 ⁶³² G/A	Cases n(%) N=130	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	26 (20)	6.8	0.68	1	1	
GA	58 (44.6)	6.3		1.13	1.26 (0.70-2.27)	0.42
AA	46 (35.3)	9.7		0.93	0.84 (0.45-1.54)	0.58
GA+AA	104 (80.0)	9.3		1.04	1.03 (0.61-1.74)	0.90
XRCC1 ²⁸⁰ G/A	Cases n(%) N=130	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	14 (10.7)	14.9	0.54	1	1	
GA	91 (70.0)	8.2		1.42	1.47 (0.71-3.05)	0.29
AA	25 (19.2)	6.3		1.27	1.31 (0.42-4.06)	0.64
GA+AA	116 (89.2)	7.5		1.39	1.62 (0.79-3.34)	0.18

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Further, the performance status KPS and ECOG were analyzed for their relation with survival and it was observed that, for the polymorphic variant *XRCC1 Arg³⁹⁹ Gln* a trend of decreased survival was observed in patients carrying single allelic variant genotype (*GA*) as KPS decreased as indicated in tables 5.59, 5.60, 5.61. The percentage of heterozygous (*GA*) subjects with KPS score between [100-80] was 53.7% as shown in (table 5.59) , 59.4% subjects had KPS score 70-60 (table 5.60) and 35% subjects were with

KPS score 50-40 (table 5.61). For patients with KPS score between [100-80] a trend of better survival (H.R.=0.78, 95%C.I.=0.50-1.20, p=0.26) was observed however, patients with KPS (70-60) had an intermediate survival (H.R.=1.23, 95%C.I.=0.61-2.45, p=0.55) table 5.60 compared to other extremes, as patients with KPS score between [50-40] had the least survival and high death rate (H.R.=2.05, 95%C.I.=0.28-14.8, p=0.47) as shown in table 5.61.

On analyzing KPS for other polymorphic sites it was observed that subjects with KPS score between [100-80] and carrying the wild (*AA*) type genotype for polymorphic site *XRCC1 Pro*²⁰⁶*Pro* a relatively better MST of 10.2 months was observed than for subjects with other genotypic combinations as shown in table 5.59. In case of *XRCC1 Arg*²⁸⁰*His* MST of 16.1 months was observed for subjects carrying wild type (*GG*) genotype and KPS score ranging from [100-80] compared to 8.6 months for heterozygous genotype (*GA*) and 12.3 months (log rank p-value=0.06) for mutant genotype (*AA*) as shown in table 5.59. Multivariate analysis for the same exhibited that subjects with single allelic variant genotype (*GA*) exhibited a trend of higher death rate (H.R.=1.65, 95%C.I.=0.82-3.33, p=0.16).

For subjects with KPS performance score ranging from [70-60] neither a significant difference was observed in the median survival time for any genetic variant nor any significance was observed for hazard rate as shown in table 5.60. Subjects carrying KPS score [50-40] and with polymorphic variant *XRCC1 Arg*¹⁹⁴ and *XRCC1 Pro*²⁰⁶*Pro* did not influence lung cancer survival. Furthermore, for subjects with polymorphic variant *XRCC1 Gln*⁶³²*Gln* single allelic variants revealed a MST of 5.2 months compared to 4.1 months for single allelic variants (Log rank p-value=0.67). Cox-regression analysis suggested a trend of very high death rate with four-fold hazard rate (H.R.=4.14, 95%C.I.=0.31-8.24, p=0.13).

Patients with KPS score between [50-40] and carrying wild type allele (*GG*) for *XRCC1 Arg*²⁸⁰*His* showcased a MST of 15.5 months as compared to 3.4 months for heterozygous genotype (*GA*) and 5.1 months for mutant genotype (*AA*) (log rank p-value=0.39). The multivariate analysis revealed that heterozygous genotype (*GA*) showed a death rate of 4 folds (H.R.=4.4, 95%C.I.=0.70-27.9, p=0.11) as indicated in table 5.61.

Table 5.59: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (KPS [100-80])

Genes	KPS [100-80]					
	Cases n (%) N=147	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value

to be cont'd on next page

Table 5.59: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (KPS [100-80]) (cont..)

Genes	KPS [100-80]					
GG	43 (29.2)	9.7	0.65		1	
GA	79 (53.7)	9.6		0.81	0.78 (0.50-1.20)	0.26
AA	25 (17.0)	9.4		0.89	0.91 (0.49-1.69)	0.77
GA+AA	104 (70.7)	9.46		0.83	0.80 (0.53-1.21)	0.31
XRCC1 Arg ¹⁹⁴ Trp	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=147	MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	119 (80.9)	9.0	0.48			
CT	27 (18.3)	12.5		0.79	0.82 (0.48-1.40)	0.47
TT	1 (0.6)	-				
CT+TT	28 (19.0)	14.0		0.76	0.80 (0.47-1.36)	0.42
XRCC1 Pro ²⁰⁶ Pro	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=147	MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	63 (42.8)	10.2	0.54		1	
AG	69 (46.9)	8.2		1.24	1.16 (0.75-1.80)	0.48
GG	15 (10.2)	9.7		1.27	1.17 (0.59-2.33)	0.65
AG+GG	84 (57.1)	9.0		1.24	1.16 (0.77-1.75)	0.46
XRCC1 Gln ⁶³² Gln	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=147	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	33 (22.4)	8.6	0.86		1	
GA	66 (44.8)	9.4		1.11	1.16 (0.69-1.96)	0.56
AA	48 (32.6)	9.6		1.01	0.97 (0.56-1.70)	0.94
GA+AA	114 (77.5)	9.66		1.07	1.11 (0.68-1.80)	0.66
XRCC1 Arg ²⁸⁰ His	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=147	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	14 (9.5)	16.1	0.06			
GA	112 (76.1)	8.6		1.76	1.65 (0.82-3.33)	0.16
AA	21 (14.2)	12.3		1.02	0.89 (0.33-2.43)	0.83
GA+AA	133 (90.4)	9.0		1.62	1.58 (0.79-3.16)	0.19

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.60: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (KPS [70-60])

Genes	KPS [70-60]					
XRCC1 Arg ³⁹⁹ Gln	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=74	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	15 (20.2)	8.8	0.12		1	
GA	44 (59.4)	4.0		1.41	1.23 (0.61-2.45)	0.55
AA	15 (20.2)	10.3		0.75	0.67 (0.27-1.68)	0.40

to be cont'd on next page

Table 5.60: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (KPS [70-60])(cont..)

Genes	KPS [70-60]					
GA+AA	59 (79.7)	4.5		1.19	1.09 (0.57-2.0)	0.78
XRCC1 Arg ¹⁹⁴ Trp	Cases n (%) N=74	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	57 (77.0)	4.5	0.76			1
CT	17 (22.9)	5.4		1.09	0.98 (0.54-1.8)	0.97
TT	0			-		
CT+TT						
XRCC1 Pro ²⁰⁶ Pro	Cases n (%) N=74	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	27 (36.4)	5.3	0.84		1	
AG	39 (52.7)	5.0		0.86	0.97 (0.55-1.70)	0.91
GG	8 (10.8)	6.5		0.88	0.70 (0.27-1.84)	0.48
AG+GG	47 (63.5)	5.0		0.86	0.91 (0.53-1.57)	0.75
XRCC1 Gln ⁶³² Gln	Cases n (%) N=74	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	13 (17)	4.3	0.30		1	
GA	32 (43.2)	4.7		0.99	1.58 (0.74-3.37)	0.23
AA	29 (39.1)	5.4		0.67	0.85 (0.39-1.82)	0.68
GA+AA	61 (82.4)	5.3		0.81	0.94 (0.48-1.83)	0.87
XRCC1 Arg ²⁸⁰ His	Cases n (%) N=74	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	8 (10)	5.0	0.70		1	
GA	51 (68.9)	5.3		0.72	0.70 (0.30-1.66)	0.43
AA	15 (20.2)	5.4		0.74	0.94 (0.27-3.19)	0.92
GA+AA	66 (89.1)	5.4		0.73	0.69 (0.30-1.57)	0.38

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.61: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (KPS [50-40])

Genes	KPS [50-40]					
XRCC1 Arg ³⁹⁹ Gln	Cases n (%) N=20	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	10 (50)	5.0	0.62		1	
GA	7 (35)	5.1		1.5	2.05 (0.28-14.8)	0.47
AA	3 (15)	9.4		0.9	-	
GA+AA	10 (50)	5.3		1.26	0.97 (0.25-3.69)	0.96
XRCC1 Arg ¹⁹⁴ Trp	Cases n (%) N=20	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value

to be cont'd on next page

Table 5.61: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (KPS [50-40]) (cont..)

Genes	KPS [50-40]					
CC	14 (70)	5.3	0.52			
CT	6 (30)	5.1		1.34	1.38 (0.41-4.60)	0.59
TT	0	-			-	
CT+TT			-			
XRCC1 Pro ²⁰⁶ Pro	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=20	MST	p-value	H.R. ^a	H.R. ^b	p-value
AA	7 (35)	5.2	0.43			
AG	6 (30)	5.1		0.70	0.80 (0.28-2.72)	0.83
GG	0	-			-	
AG+GG	-	-			-	
XRCC1 Gln ⁶³² Gln	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=20	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	1(5)	4.1	0.67		1	
GA	11 (55)	5.2		0.43	4.14 (0.31-8.24)	0.135
AA	8 (40)	5.6		0.55	-	-
GA+AA	19 (95)	5.2		0.48	-	-
XRCC1 Arg ²⁸⁰ His	Cases n (%)	Univariate analysis			Multivariate analysis	
	N=20	MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	4 (20)	15.5	0.39		1	
GA	12 (60)	3.4		1.93	4.4 (0.70-27.9)	0.11
AA	4 (20)	5.1		2.39	-	-
GA+AA	16 (80)	5.1		2.03	3.69 (0.69-19.78)	0.12

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

ECOG and polymorphic *XRCC1*

On analyzing ECOG, and its association with the polymorphic variant of *XRCC1* gene it was demonstrated that patients who had ECOG score of [0-1] and carrying single allelic variant (*GA*) for *XRCC1 Arg³⁹⁹Gln* depicted a better survival (H.R.=0.65, 95%C.I.=0.40-1.0, p=0.07) as shown in table 5.62. The next SNP *XRCC1 Arg¹⁹⁴Trp* was observed not to have any association with survival of lung cancer subjects. In case of *XRCC1 Pro²⁰⁶Pro* patients with mutant genotype (*GG*) and ECOG score between [0-1] were representing a trend exactly opposite to that observed in *XRCC1 Arg³⁹⁹Gln*. Patients had a shorter survival (MST=9.0 vs 11.7) (H.R.=1.43, 95%C.I.=0.66-3.99, p=0.36) for mutant subjects for polymorphic *XRCC1 Pro²⁰⁶Pro*. The next SNP *XRCC1 Gln⁶³²Gln* did not modulate survival of lung cancer subjects as illustrated in table 5.62. *XRCC1 Arg²⁸⁰His* when examined, subjects with ECOG score 0-1 and carrying wild type genotype showed a

MST of 15.9 months compared to 12.3 months for double allelic variant genotype (*AA*) (log rank p-value=0.12). Cox regression analysis exhibited a decrease in hazards rate (H.R.=0.57, 95%C.I.=0.19-1.67, p=0.31).

Furthermore, for *XRCC1 Arg³⁹⁹Gln* patients with ECOG score of [2] and harboring heterozygous (*GA*) genotype had an lower MST as compared to wild (*GG*) type subjects (MST=4.5 vs 9.3, log rank p-value=0.43), Cox-regression analysis revealed an increased death rate (H.R.=1.63, 95%C.I.=0.78-3.37, p=0.18) for subjects with heterozygous genotype and ECOG score of [2]. Polymorphic variant *XRCC1 Arg¹⁹⁴Trp* and *XRCC1 Pro²⁰⁶Pro* did not show any association with survival of lung cancer subjects as table 5.63 indicates in relation to ECOG statistics. The median survival time for wild genotype (*GG*) for subjects with genetic variant *XRCC1 Gln⁶³²Gln* was 12.4 months compared to 4.7 months for single allelic variants (*GA*). A trend towards increased hazard rate was observed for heterozygous subjects (H.R.=1.8, 95%C.I.=0.76-4.27, p=0.17). Trend of increase in hazard rate amplified further with increased in variant allele for subjects with ECOG score [2]. Mutant (*AA*) subjects had a death rate of 2.3 fold (H.R.=2.33 , 95%C.I.=0.78-6.97, p=0.12) which decreased for combined variants (*GA+AA*) (H.R.=2.0, 95%C.I.=0.88-4.57, **p=0.093**). Double allelic variants (*AA*)for the polymorphic variant *XRCC1 Arg²⁸⁰His* showcased a high death rate (H.R.=1.61, 95%C.I.=0.32-8.10, p=0.55).

As shown in table 5.64 lung cancer subjects with ECOG score between [3-4] and carrying mutant genotype (*AA*) for *XRCC1 Arg³⁹⁹Gln* exhibited a significant prolonged survival as compared to patients with wild type genotype (MST=9.4 vs 4.3, log rank p-value=0.44). Cox-regression analysis showed that death rate was 0.24 fold (H.R.=0.24, 95%C.I.=0.06-0.96, **p=0.04**). *XRCC1 Arg¹⁹⁴Trp* like other two ECOG score groups did not show any influence on survival of lung cancer subjects as shown in table 5.64.

Furthermore, subjects with mutant genotype (*GG*) of polymorphic variant *XRCC1 Pro²⁰⁶Pro* with ECOG score [3-4] had a high hazard ratio (H.R.=1.66, 95%C.I.=0.06-45.9, p=0.76) as MST was less compared to patients with wild type genotype and same ECOG score (MST=3.4 vs 5.0, log rank p-value=0.82). For the next SNP *i.e.* *XRCC1 Gln⁶³²Gln* it was studied that carriers of wild type genotype (*GG*) had a better MST of 4.1 months compared to median survival time of 2.1 months for heterozygous (*GA*)subjects. Double allelic variants (*AA*) with ECOG 3-4 had a better MST of 6.2 months (log rank p-value=0.33) and a significant better survival rate (H.R.=0.17, 95%C.I.=0.03-0.90, **p=0.03**).

Table 5.62: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (ECOG [0-1])

Genes	Cases n(%) N=127	ECOG (0-1)				
		Univariate analysis			Multivariate analysis	
XRCC1 Arg ³⁹⁹ Gln		MST	p-value	H.R ^a	H.R ^b	p-value
GG	39 (30.7)	9.7	0.44		1	
GA	63 (49.6)	10.2		0.74	0.65 (0.40-1.00)	0.07
AA	25 (19.6)	11.26		0.78	0.67 (0.34-1.30)	0.24
GA+AA	88 (69.2)	10.6		0.76	0.68 (0.43-1.07)	0.10
XRCC1 Arg ¹⁹⁴ Trp	Cases n(%) N=127	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
CC	101 (79.5)	9.4	0.5		1	
CT	25 (19.6)	12.5		0.79	0.79 (0.45-1.39)	0.42
TT	1 (0.7)				-	
CT+TT	26 (20.4)	12.56		0.77	0.77 (0.44-1.36)	0.38
XRCC1 Pro ²⁰⁶ Pro	Cases n(%) N=127	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
AA	57 (44.8)	11.7	0.74		1	
AG	58 (45.6)	10.1		1.08	1.05 (0.67-1.66)	0.81
GG	12 (9.4)	9.0		1.31	1.43 (0.66-3.99)	0.36
AG+GG	70 (55.1)	10.0		1.12	1.09 (0.71-1.69)	0.66
XRCC1 Gln ⁶³² Gln	Cases n(%) N=127	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	27 (21)	8.2	0.79			
GA	59 (46.4)	10.6		0.94	0.91 (0.51-1.59)	0.74
AA	41 (32.2)	11.7		0.83	0.81 (0.43-1.50)	0.51
GA+AA	100 (78.7)	10.6		0.89	0.87 (0.51-1.48)	0.62
XRCC1 Arg ²⁸⁰ His	Cases n(%) N=127	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	15 (11)	15.9	0.12			
GA	93 (73.2)	9.4		1.43	1.51 (0.77-2.95)	0.22
AA	19 (14.9)	12.3		0.77	0.57 (0.19-1.67)	0.31
GA+AA	112 (88.1)	9.8		1.30	1.36 (0.70-2.65)	0.36

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.63: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (ECOG [2])

Genes	ECOG (2)					
	Cases n(%) N=73	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
XRCC1 Arg ³⁹⁹ Gln						
GG	14 (19.1)	9.3	0.43		1	
GA	47 (64.3)	4.5		1.53	1.63 (0.78-3.37)	0.18
AA	12 (16.4)	5.4		1.43	1.64 (0.58-4.63)	0.35
GA+AA	59 (80.8)	5.0		1.51	1.67 (0.83-3.34)	0.14
XRCC1 Arg ¹⁹⁴ Trp						
	Cases n(%) N=73	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
CC	61(83.5)	5.4	0.76		1	
CT	12 (16.4)	5.0		1.18	1.15(0.57-2.31)	0.68
TT	0			-		
CT+TT	-	-			-	
XRCC1 Pro ²⁰⁶ Pro						
	Cases n(%) N=73	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
AA	23 (31.5)	5.4	0.84		1	
AG	40 (54.7)	4.5		0.94	1.17 (0.62-2.21)	0.62
GG	10 (13.6)	8.8		0.78	0.69 (0.26-1.82)	0.46
AG+GG	50 (68.4)	5.0		0.90	0.97 (0.54-1.74)	0.92
XRCC1 Gln ⁶³² Gln						
	Cases n(%) N=73	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	10 (13)	12.4	0.19		1	
GA	32 (43.8)	4.7		2.01	1.8 (0.76-4.27)	0.17
AA	31 (42.4)	5.4		1.75	2.33(0.78-6.97)	0.12
GA+AA	63 (86.3)	5.0		1.88	2.0 (0.88-4.57)	0.093
XRCC1 Arg ²⁸⁰ His						
	Cases n(%) N=73	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	8 (10)	5.9	0.98		1	
GA	55 (75.3)	5.3		0.96	0.81 (0.32-2.00)	0.65
AA	10 (13.6)	6.3		1.02	1.61 (0.32-8.10)	0.55
GA+AA	65 (89.0)	5.3		0.97	0.90 (0.37-2.19)	0.82

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.64: Genetic determinants of *XRCC1* and their association with clinical parameters and outcomes (ECOG 3-4)

Genes	ECOG 3-4					
	Cases n(%) N=41	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
XRCC1 Arg ³⁹⁹ Gln						
GG	15 (36.5)	4.3	0.44		1	
GA	20 (48.7)	2.1		1.02	1.21 (0.46-3.15)	0.69
AA	6 (14.6)	9.4		0.55	0.24 (0.06-0.96)	0.04
GA+AA	26 (63.4)	5.1		0.86	0.85 (0.36-2.02)	0.72
XRCC1 Arg ¹⁹⁴ Trp						
	Cases n(%) N=41	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
CC	28 (68.2)	4.1	0.81		1	
CT	13 (31.7)	5.2		0.92	0.78 (0.35-1.74)	0.56
TT	0	-			-	
CT+TT		-		-	-	
XRCC1 Pro ²⁰⁶ Pro						
	Cases n(%) N=41	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
AA	17 (41.4)	5.0	0.82		1	
AG	23 (56.0)	4.3		1.07	1.05 (0.53-2.08)	0.88
GG	1 (2.4)	3.4		1.86	1.66 (0.06-45.9)	0.76
AG+GG	24 (58.5)	4.3		1.09	1.04 (0.53-2.07)	0.89
XRCC1 Gln ⁶³² Gln						
	Cases n(%) N=41	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	10 (24)	4.1	0.33		1	
GA	18 (43.9)	2.1		0.98	0.93 (0.30-2.85)	0.90
AA	13 (31.7)	6.2		0.59	0.17 (0.03-0.90)	0.03
GA+AA	31 (75.6)	5.1		0.77	0.66 (0.24-1.80)	0.42
XRCC1 Arg ²⁸⁰ His						
	Cases n(%) N=41	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
GG	3 (7.3)	14.9	0.80		1	
GA	27 (65.8)	3.4		1.48	1.14 (0.29-4.46)	0.84
AA	11 (26.8)	5.2		1.49		
GA+AA	138 (92.6)	4.3		1.48	1.54 (0.42-5.62)	0.51

^a-Unadjusted Hazards ratio for Kaplan meier analysis, ^b-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.10.7 Combinatorial effect of double combinations of *XRCC1* SNPs on overall survival of lung cancer patients

The cumulative effect of the five SNPs on overall survival was also analyzed in our study both in combinations of doublets and triplets. The genotypic frequency distribution and

survival of five polymorphic variants are enlisted in table 5.65. Overall none of the combinations could significantly predict a prominent increase or decrease in the survival rate of the study population. However, some trends were visualized from the analysis. Genotypic combination of *XRCC1 399* and *194* had the highest representation frequency (63.0%) and showed that the patients having both the wild genotype (*GG+CC*) for the above combination had the highest (MST=8.8) as compared to heterozygous genotypes (MST=5.9) however, there was no significant difference between the time of survival ($p=0.69$) as shown in table 5.65. For the genotypic combination of *XRCC1 399* and *206* 22.5% of patients were carrying alleles of wild type genotype for both SNPs i.e. (*GG+AA*), whereas 73.1% subjects were heterozygous for both SNPs (*GA+AG*) and 4.3% subjects had both mutant genotype (*AA+GG*). Subjects carrying the mutant genotypic (*AA+GG*) group showed a visible trend of the highest median survival time (MST=8.8) when compared with the wild (*GG+AA*) (MST=5.4) and heterozygous (*GA+AG*) (MST=5.8) genotype.

Genotypic combination of *XRCC1 399* and *632* for mutant carriers (*AA/AA*) depicted the longest survival with MST=9.4 and low univariate death rate (H.R.=0.6). When H.R. was adjusted with factors like age, sex, smoking hazard rate reduced further (H.R.=0.42, 95%C.I.=0.16-1.31, $p=0.08$). For the genotypic combination of *XRCC1 399* and *280* 65% of lung cancer subjects were carrying the wild type (*GG+GG*) allele for both the above SNPs and sequentially the highest MST of 16.1 months as compared to subjects with heterozygous genotype (MST=6.4). Single allelic (*GA/GA*) carriers for the combination had a higher representation compared to double allelic variants (*AA/AA*) (87.1% vs 5.6%). Regression analysis for single allelic carriers revealed a high hazard rate (H.R.=2.09, 95%C.I.=0.85-5.16, $p=0.08$).

The next genotypic combination analyzed in the sequence is of *XRCC1 194* with other SNPs such as *206*, *632*, *280* as shown in table 5.65. However, the genotypic combinations of *XRCC1 206* and *632* did not correlate with overall survival as shown in table 5.65. However, the genotypic combination of *XRCC1 194* and *280* showed that the subjects carrying wild type genotype (*CC+GG*) had a MST of 12.0 months compared to 5.3 months of heterozygous genotype (*CT+AG*). The multivariate analysis showed a visible trend of increase in death rate (H.R.=1.45, 95%C.I.=0.69-3.05, $p=0.32$). Further, the combination of *XRCC1 206* with *280* for heterozygous genotype (*AG/GA*) depicted a shorter MST of 6.1 months with a high 2.0 fold death rate (H.R.=2.0, 95%C.I.=0.71-5.9, $p=0.18$).

Table 5.65: Genotypic combinations and overall survival

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R. ^a	H.R. ^b	p-value	FDR _p
XRCC1 Arg399Gln, G/A & XRCC1 Arg194Trp, C/T	N=84						
GG/CC	53 (63.0)	8.8	0.69	1	1		
GA/CT	30 (35.7)	5.9		1.00	1.22 (0.68-2.16)	0.49	0.83
AA/TT	1 (1.19)	-		-			
GA+AA/CT+T	31 (36.9)	5.9		0.97	1.20 (0.68-2.13)	0.52	0.88
XRCC1 Arg399Gln, G/A & XRCC1 Pro206Pro, A/G	N=93						
GG/AA	21 (22.5)	5.4	0.90	1	1		
GA/AG	68 (73.1)	5.8		0.90	0.94 (0.53-1.65)	0.83	0.83
AA/GG	4 (4.3)	8.8		0.80	4.21 (0.86-20.5)	0.07	0.08
GA+AA/AG+GG	72 (77.4)	5.8		0.89	0.91 (0.52-1.58)	0.74	0.88
XRCC1 Arg399Gln, G/A & XRCC1 Gln632Gln, G/A	N=94						
GG/GG	18 (19.1)	5.1	0.15	1	1		
GA/GA	54 (57.4)	5.2		1.04	1.17 (0.59-2.31)	0.63	0.83
AA/AA	22 (23.4)	9.4		0.6	0.42 (0.16-1.31)	0.08	0.08
GA+AA/GA+AA	76 (80.8)	5.9		0.88	0.89 (0.47-1.69)	0.73	0.88
XRCC1 Arg399Gln, G/A & XRCC1 Arg280Gln, G/A	N=107						
GG/GG	7 (6.5)	16.1	0.37	1	1		
GA/GA	94 (87.8)	6.4		1.71	2.09 (0.85-5.16)	0.72	
AA/AA	6 (5.6)	6.6		1.28	-	-	-
GA+AA/GA+AA	100 (93.4)	6.0		1.68	1.93 (0.79-4.69)	0.14	0.7
XRCC1 Arg194Try, C/T & XRCC1 Pro206 Pro, A/G	N=87						
CC/AA	68 (78.1)	6.7	0.52	1	1		
CT/AG	18 (20.6)	5.9		1.21	1.09 (0.60-1.99)	0.77	0.83
TT/AA	1 (1.1)	-		-	-		
CT+TT/AG+AA	19 (21.8)	6.1		1.14	1.04 (0.57-1.90)	0.88	0.88

to be cont'd on next page

Table 5.65: Genotypic combinations and overall survival (cont..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R. ^a	H.R. ^b	p-value	FDR _p
XRCC1 Arg194Try, C/T & XRCC1 Gln632Gln, G/A	N=61						
CC/GG	42 (68.8)	6.8	0.57	1	1		
CT/GA	19 (31.1)	5.3		1.18	1.17 (0.62-2.23)	0.62	0.83
XRCC1 Arg194Try, C/T & XRCC1 Arg280His, G/A	N=52						
CC/GG	18 (34.6)	12.0	0.46	1	1		
CT/GA	34 (65.3)	5.3		1.26	1.45 (0.69-3.05)	0.32	0.83
XRCC1 Pro206Pro, A/G & XRCC1 Gln632Gln, G/A	N=92						
AA/GG	9 (9.7)	4.1	0.81	1	1		
AG/GA	77 (83.6)	7.2		1.29	0.80 (0.29-2.17)	0.66	0.83
AA/GG	6 (6.5)	8.9		1.35			
AG+AA/GA+GG	83 (90.2)	7.2		1.29	0.83 (0.30-2.2 5)	0.72	0.88
XRCC1 Pro206Pro, A/G & XRCC1 Arg280His, G/A	N=91						
AA/GG	7 (7.6)	12.0	0.087	1	1		
AG/GA	81 (89.0)	6.1		2.30	2.0 (0.71-5.9)	0.18	0.81
GG/AA	3 (3.2)	12.3		1.08	-		
AG+GG/GA+AA	84 (92.3)	6.8		2.26	1.97 (0.68-5.7)	0.20	0.7

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.10.8 Combined effect of triple combinations of *XRCC1* SNPs on overall survival

The study was further comprehended for the triple combinations of the five polymorphic variants of *XRCC1*. The triple combinations *XRCC1* 399, 194, 206, *XRCC1* 399, 206, 632 were observed not to modify lung cancer survival. For the triple combination of *XRCC1* 399, 206 and 280 subjects carrying heterozygous genotype (GA/AG/GA) showed a MST of 5.4 months and a very high death rate of 3.4 fold (H.R.=3.41, 95% C.I.=0.49-31.3, p=0.19). The triple combination of *XRCC1* 399, 194, 280, *XRCC1* 399, 194, 632, 194, 206, 280 did not modulate lung cancer risk as shown in table 5.66. The next triple combination of *XRCC1* 194, 206, 632 showed a MST of 6.1 months for heterozygous genotype compared to 4.0 months for wild type genotype analysis and showed a very small death rate (H.R.=0.25, 95% C.I.=0.03-2.10, p=0.20) as shown in table 5.66.

Table 5.66: Genotypic combinations and overall survival

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R. ^a	H.R. ^b	p-value	FDR _p
XRCC1Arg 399 Gln, G/A & XRCC1Arg194Trp, C/T & XRCC1Pro206Pro, A/G	N=22						
GG/CC/AA	15 (68.1)	5.4	0.53	1	1		
GA/CT/AG	6 (27.2)	5.0		1.09	0.5 (0.09-3.64)	0.55	0.79
GA+AA/CT+TT/AG+GG	7 (31.8)	5.5		0.78	0.45 (0.082-2.50)	0.36	0.88
XRCC1 Arg399Gln, G/A & XRCC1 Pro206Pro, A/G & XRCC1 Gln632Gln, G/A	N=45						
GG/AA/GG	4 (8.8)	3.7	0.11	1	1		
GA/AG/GA	40 (88.8)	5.1		0.45	0.88 (0.31-3.00)	0.85	0.85
GA+AA/AG+GG/GA+AA	41 (91.1)	5.1		0.46	0.86 (0.23-3.15)	0.82	0.88
XRCC1 Arg399Gln, G/A & XRCC1 Pro206Pro, A/G & XRCC1 Arg280His, G/A	N=47						
GG/AA/GG	2 (4.2)	-	0.19	1	1		
GA/AG/GA	45 (95.7)	5.4		3.41	3.44 (0.49-31.3)	0.19	0.70
XRCC1 Arg399Gln, G/A & XRCC1 Arg194Trp, C/T & XRCC1 Arg280His, G/A	N=24						
GG/CC/GG	4 (16.6)	14.1	0.93	1	1		
GA/CT/GA	20 (83.3)	3.9		1.04	1.29 (0.20-8.0)	0.78	0.85
XRCC1 Arg399Gln, G/A & XRCC1 Arg194Trp, C/T & XRCC1 Gln632Gln, G/A	N=27						
GG/CC/GG	16 (59.2)	5.13	0.93	1	1		
GA/CT/GA	11 (40.7)	5.2		0.96	0.67 (0.25-1.80)	0.44	0.79
XRCC1 Arg194Trp, C/T & XRCC1 Pro206Pro, A/G & XRCC1 Gln632Gln, G/A	N=19						

to be cont'd on next page

Table 5.66: Genotypic combinations and overall survival (cont..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R. ^a	H.R. ^b	p-value	FDR _p
CC/AA/GG	8 (42.1)	4.0	0.95	1	1		
CT/AG/GA	11 (57.8)	6.1		0.97	0.25 (0.03-2.10)	0.20	0.70
XRCC1 Arg194Trp, C/T & XRCC1 Pro206Pro, A/G & XRCC1 Arg280His, G/A	N=15						
CC/AA/GG	5 (33.3)	12.0	0.21	1	1		
CT/AG/GA	10 (66.6)	6.1		2.11	0.49 (0.045-5.4)	0.57	0.79

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.10.9 Relationship of *XRCC1* polymorphism and their correlation with objective response after chemotherapy

Associations between polymorphisms and the efficacy outcome of response rate and clinical benefits were studied. The patients were divided into good and bad responders on the basis of RECIST criteria, where patient were classified into four groups, CR (complete response), PR (partial response), SD (stable disease) and PD (progressive disease). The adjusted odds ratio was calculated for each polymorphic site taking the wild genotype as a reference and age, gender, smoking, stage, histology, KPS, ECOG and chemotherapeutic regimen were the covariates. The objective response data was available for 130 patients. Seventy-three patients experienced a good response (CR+PR) whereas fifty-seven patients (SD+PD) did not respond to platinum based chemotherapy. We report that none of the five SNPS of the *XRCC1* gene could be clearly classified either as good or bad responders and act as significant predictors of response rate or clinical benefit overall.

In case of *XRCC1 Arg³⁹⁹Gln* the frequency of subjects with heterozygous (*GA*) genotype was 57.8% in poor responders and 52.0% in good responders as shown in table 5.67. However, logistic regression analysis could not classify the polymorphic variant either as good or poor responder. Similarly, a slightly higher percentage of subjects with heterozygous (*AG*) genotype for *XRCC1 Pro²⁰⁶Pro* was observed for poor responders as compared to good responders (54.3% vs 46.5%). The percentage distribution of combined variants genotype were almost similar in both group of responders and non-responders as a result none of the genotype of *XRCC1 Arg³⁹⁹Gln*, *XRCC1 Gln⁶³²Gln*, *XRCC1 Arg²⁸⁰His* could act as a significant predictor of chemotherapeutic response rate.

Table 5.67: Effect of *XRCC1* genetic variants on treatment outcome

Genotype	Good response CR+PR n(%) N = 73	Poor response SD+PD n (%) N = 57	AOR (95% C.I.) ^a	p ^b
<i>XRCC1 Arg³⁹⁹Gln G/A</i>				
GG	20 (27.3)	15 (26.3)	1	
GA	38 (52.0)	33 (57.8)	0.74 (0.3-1.85)	0.53
AA	15 (20.5)	9 (15.7)	1.28 (0.37-4.3)	0.69
GA+AA	53 (72.6)	42 (73.6)	0.92 (0.39-2.15)	0.85
<i>XRCC1 Arg¹⁹⁴Trp C/T</i>				
CC	60 (82.1)	48 (85.9)	1	
CT	13 (17.8)	8 (14.0)	0.90 (0.30-2.67)	0.85
TT	-	-	-	-

to be cont'd on next page

Table 5.67: Effect of *XRCC1* genetic variants on treatment outcome

Genotype	Good response CR+PR n (%) N = 73	Poor response SD+PD n (%) N = 57	AOR (95% C.I.) ^b	p ^a
CT+TT	-	-	-	-
XRCC1 Pro²⁰⁶Pro A/G				
AA	33 (45.2)	23 (40.3)	1	
AG	34 (46.5)	31 (54.3)	0.74 (0.33-1.69)	0.48
GG	6 (8.2)	3 (5.2)	1.0 (0.18-5.7)	0.96
AG+GG	40 (54.7)	34 (59.6)	0.77 (0.35-1.69)	0.52
XRCC1 Gln⁶³²Gln G/A				
GG	16 (21.9)	12 (40.3)	1	
GA	29 (39.7)	30 (52.6)	0.65 (0.22-1.8)	0.42
AA	28 (38.3)	15 (26.3)	1.29 (0.43-3.83)	0.64
GA+AA	57(78.0)	45 (78.9)	0.85 (0.33-2.19)	0.74
XRCC1 Arg²⁸⁰His G/A				
GG	8 (10.9)	6 (10.5)	1	
GA	55 (75.3)	35 (61.4)	1.33 (0.36-4.8)	0.65
AA	10 (13.6)	16 (28.0)	0.32 (0.03-2.87)	0.32
GA+AA	65 (89.0)	51 (89.4)	1.11 (0.31-3.9)	0.86

(a)-Adjusted Odds ratio, (b)- 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, histology, stage, KPS, ECOG. Abbreviations: CR-Complete reponse, PR-Partial reponse, SD-Stable disease, PD-Progressive disease, AOR-Adjusted odds ratio, C.I.-Confidence interval

5.11 Survival analysis of DSBR genes *XRCC3*

5.11.1 Clinical characteristics of lung cancer subjects with polymorphic *XRCC3* genotype

The distribution of characteristic clinical features and response to platinum based chemotherapy of 325 study subjects are tabulated in table 5.68. The study recorded 85.5% (278) events or deaths with a median follow up time of 11.5 months. The average age for the study subject was 58.9±10.21. The study subject included 86.7% (282) males and 13.2% (43) females. Smoking status for the study was defined with 83.3% (271) smokers and 16.6% (54) non-smokers. Clinical specifications such as histology, stage, TNM were recorded from clinical history of the patients. Histological analysis from samples

of these patients showed that there were 32% (104) adenocarcinoma , 38.7% (126) squamous cell carcinoma, 27.6% (90) small cell carcinoma, and 1.5% (5) were unclassified. The clinical staging distinguished study subjects majorally in stage III (168; 51.6%) and IV (141; 43.3%) as there were very less patients with stage I (2;0.6) and II (8;2.4), however there were 16 patients with unknown stage. TNM data was available for 310 patients as shown in table 5.68. Tumor extension for the study group was classified as Tx,T1,T2,T3,T4,unknown tumor extension, the respective frequency distribution was as 2.4% (8), 4.6% (15), 21.6% (41), 24% (78), 51.6% (168) and 4.6% (15) respectively. Lymph node distribution revealed that there were 12.0% (39) subjects within N0 category, 10.1% (33) within N1, 45.5% (148) with N2 classification, 27.6% (90) under N3 category and 4.6% (15) with unknown lymph node invasion. Metastatic stage defined there were 54.1% (176) M0 without metastasis and 41.2% (134) were with distant metastasis, and 4.6% (15) with unknown metastatic state. The chemotherapeutic regimen administered showed there were 36.7% (83) subjects who were given cisplatin/carboplatin + docetaxel, 33.1% (75) were given cisplatin/carboplatin+ irinotecan and 30.0% (30.0) were given cisplatin/carboplatin + pemetrexed.

Table 5.68: Patients Characteristics and clinical features of *XRCC3*

Patients characteristics	No.of Patients, n(%) N=325
Age(Years)	
Mean±SD Range	58.9±10.21 (28-83)
Gender	
Male	282 (86.7)
Female	43 (13.2)
Smoking status	
Smokers	271 (83.3)
Non-smokers	54 (16.6)
Histology	
ADCC	104 (32)
SQCC	126 (38.7)
SCLC	90 (27.6)
Others	5 (1.5)
TNM staging	
I	2 (0.6)
II	8 (2.4)
III	168 (51.6)
IV	141 (43.3)
Unclassified	6 (1.8)
ECOG performance status	
0-1	152 (46.7)
2=<	164 (50.4)

to be cont'd on next page

Table 5.68: Patients Characteristics and clinical features of *XRCC3*

Patients characteristics	No.of Patients, n(%) N=325
Others	9 (2.7)
KPS performance status	
100-80	201 (61.8)
<70	115 (35.3)
Others	9 (2.7)
Tumor size	
Tx	8 (2.4)
T1	15 (4.6)
T2	41 (12.6)
T3	78 (24)
T4	168 (51.6)
Others	15 (4.6)
Lymph node	
N0	39 (12.0)
N1	33 (10.1)
N2	148 (45.5)
N3	90 (27.6)
Others	15 (4.6)
Metastasis	
M0	176 (54.1)
M1	134 (41.2)
Others	15 (4.6)
Median Follow up (months)	11.05
Events, Death	278 (85.5)
Chemotherapy Regimen	N =226
Docetaxel + Cisplatin/Docetaxel + Carboplatin	83 (36.7)
Irinotecan + Cisplatin/Irinotecan + Carboplatin	75 (33.1)
Pemetrexed + Cisplatin/Pemetrexed + Carboplatin	68 (30.0)

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects, (a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

5.11.2 Polymorphic *XRCC3 T241M* and Survival- Overall and histology based

The genotypic distribution and overall survival of the study subject is specified in table 5.69 . Statistical analysis for the single nucleotide polymorphism (SNP) *XRCC3*

*Thr*²⁴¹*Met* reported there were 278 patients dead during the study and 47 alive. No affect was observed on survival in relation to any genotypic combination of the SNP. The frequency distribution suggested there were 49.3% patients carrying wild type genotype (*CC*) and 50.6% patients carrying heterozygous subjects (*CT*). Subjects with wild type genotype (*CC*) showed a MST 8.8 months compared to 8.0 months for heterozygotes (*CT*) (Log rank p-value=0.81). Accordingly, after multivariate analysis, no effect was observed on hazard rate (H.R.=1.1, 95%C.I.=0.85-1.41, p=0.45). Prognostic effect was assessed further according to the histological subtypes where again no effect was observed for any histology. However, in case of SCLC MST for wild type (*CC*) was 9.0 months compared to the heterozygous subjects (*CT*) 1.7 months. The difference between the months was not significant(Log rank p-value=0.73) as shown in table 5.69.

Table 5.69: *XRCC 3* polymorphic variant and their association with overall survival

Genes	Cases N=325 (n%)	Univariate analysis			Multivariate analysis	
		MST (months)	Log rank p- value	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	p-value
XRCC3 Thr²⁴¹Met C>T						
CC	160 (49.3)	8.8	0.81	1	1	
CT	164 (50.6)	8.0		1.02 (0.81-1.30)	1.1 (0.85-1.41)	0.45
TT	1	-	-	-	-	-
ADCC	N=104					
CC	49 (47.1)	8.8	0.51	1	1	
CT	55 (52.8)	9.0		1.15 (0.75-1.74)	1.16 (0.74-1.82)	0.51
SQCC	N= 126					
CC	50 (39.6)	6.4	0.8	1	1	
CT	75(59.5)	8.0		1.04 (0.70-1.54)	1.06 (0.7-1.61)	0.75
TT	1	-	-	-	-	-
SCLC	N=90					
CC	60 (66.6)	9.0	0.73	1	1	
CT	30 (33.3)	7.1		0.92 (0.58-1.45)	1.1 (0.67-1.8)	0.65

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.11.3 *XRCC 3* polymorphism and overall survival based upon chemotherapy regimens

The three chemotherapeutic regimens *i.e.* cisplatin /carboplatin + docetaxel, cisplatin/ carboplatin + irinotecan, cisplatin/ carboplatin+ pemetrexed were given to the 226 study subjects accordingly. 83 (36.7%) patients were administered cisplatin/carboplatin + docetaxel, 75 (33.1%) were given cisplatin/carboplatin + irinotecan, and remaining 68 (30.0%) were given cisplatin/carboplatin+ pemetrexed. None of the regimens independently revealed any effect on survival, however as shown in table 5.70 for subjects treated with cisplatin/carboplatin + docetaxel a trend was observed where heterozygous (*CT*) subjects exhibited a better MST compared to subjects with wild type (*CC*) genotype (MST= 9.7 vs 6.7, Log rank p-value=0.83). However, for the rest of the two regimens *i.e.* cisplatin/carboplatin + irinotecan and cisplatin/carboplatin+ pemetrexed lung cancer patients with wild type genotype (*CC*) had a better survival compared to heterozygous subjects (*CT*) as shown in tables 5.71 and 5.72.

Table 5.70: Association of *XRCC3* genetic variant with survival of lung cancer patients treated with chemotherapy regimen (Cisplatin/Carboplatin+Docetaxel)

Genes	Cisplatin/Carboplatin+Docetaxel					
	Cases n(%)	Univariate analysis			Multivariate analysis	
C/T		MST	p-value	H.R ^a	H.R ^b	p-value
CC	33 (39.7)	6.7			1	
CT	50 (60.2)	9.7	0.83	0.94 (0.58-1.54)	1.08 (0.64-1.84)	0.75

(a)Unadjusted Hazards ratio for Kaplan meier analysis, (b) Hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.71: Association of *XRCC3* genetic variant with survival of lung cancer patients treated with chemotherapy regimens (Cisplatin/Carboplatin+Irinotecan) and Overall Survival

Genes	Cisplatin/Carboplatin+Irinotecan					
	Cases n(%)	Univariate analysis			Multivariate analysis	
C/T		MST	p-value	H.R ^a	H.R ^b	p-value
CC	42 (56.0)	10.4		1	1	
CT	33 (44.0)	7.1	0.89	1.03 (0.63-1.68)	1.19 (0.69-2.07)	0.52

(a) Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.72: Association of *XRCC3* genetic variant with survival of lung cancer patients treated with chemotherapy regimens (Cisplatin/Carboplatin+Pemetrexed)

Genes	Cisplatin/Carboplatin+Pemetrexed					
	Cases n(%) N=68	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R ^a	H.R ^b	p-value
CC	29 (42.6)	10.6		1	1	
CT	39 (57.3)	9.2	0.49	1.20 (0.79-2.30)	1.07 (0.58-1.95)	0.82

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.11.4 Role of *XRCC3* polymorphism and its outcome with objective response

Study subjects were classified as good responders and poor responders depending on response towards drugs. Subjects having partial response (PR) and complete response (CR) together were classified as "Good responders" and subjects with stable disease and progressive disease together were categorized as "Poor responders". The sample size for the objective response for polymorphic *XRCC3 Thr²⁴¹Met* was 186. The frequency distribution suggested there was a similar distribution of subjects under good and poor responders for both wild type subjects (47.0 vs 47.6%) and single allelic variants (52.9 vs 52.3%). Regression analysis revealed a OR of 1.5 fold (OR=1.5, 95%C.I.=0.74-3.05, p=0.24) indicating that heterozygous subjects *CT* hints towards a trend where it could be classified as good responders as shown in table 5.73

Table 5.73: Association of *XRCC3* genetic variant with chemotherapy response

Genotype	Good response CR+PR n(%) N = 102	Poor response SD+PD n (%) N = 84	AOR (95% C.I.) ^a	p ^b
<i>XRCC3 Thr²⁴¹ Met</i>				
CC	48 (47.0)	40 (47.6)	1	
CT	54 (52.9)	44 (52.3)	1.51 (0.74-3.05)	0.24

(a)-Adjusted Odds ratio, (b)- 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, histology, stage, KPS, ECOG. Abbreviations: CR-Complete reponse, PR-Partial reponse, SD-Stable disease, PD-Progressive disease, AOR-Adjusted odds ratio, C.I.-Confidence interval

5.12 Survival analysis of DSBR gene *XRCC4*

5.12.1 Clinical characteristics of lung cancer patients with polymorphic *XRCC4*

Survival analysis was carried out for *XRCC4* intron 3. The clinical characteristics are detailed in table 5.74. It was observed the mean age for the study subjects was 57.9 ± 10.6 . There were 86.9% deaths reported in the present analysis. The analysis reported there were 86.5 % (213) males and 13.4% (33) females. Out of 246 subjects there were 79.6% smokers and 20.3% non-smokers. Pathological analysis revealed there were 37.3% (92) adenocarcinomas, 38.2% (94) squamous cell carcinoma subjects and 23.9% (59) small cell carcinoma subjects. Classification according to stage classified the subjects into stage III and stage IV subjects. There were 0.8 % (2) stage I, 3.2% (8) stage II, 48.3% (119) stage III and 44.3% (103) stage IV subjects.

Table 5.74: Patients Characteristics and clinical features of *XRCC4*

Patients characteristics	No.of Patients, n(%) N=246
Age(Years)	
Mean $\pm$ SD Range	57.9 $\pm$ 10.6 (28-86)
Gender	
Male	213 (86.5)
Female	33 (13.4)
Smoking status	
Smokers	196 (79.6)
Non-smokers	50 (20.3)
Histology	
ADCC	92 (37.3)
SQCC	94 (38.2)
SCLC	59 (23.9)
Others	1 (0.4)
TNM staging	
I	2 (0.8)
II	8 (3.2)
III	119 (48.3)
IV	109 (44.3)
Unclassified	8 (3.2)
ECOG performance status	
0-1	214 (86.9)
2= $\leq$	31 (12.6)
Others	1 (0.4)

to be cont'd on next page

Table 5.74: Patients Characteristics and clinical features of *XRCC4*

KPS performance status	
100-80	155 (63.0)
<70	81 (32.9)
Others	10 (4.0)
Events, Death	214 (86.9)
Chemotherapy Regimen	N=127
Docetaxel + Cisplatin/Docetaxel + Carboplatin	45 (35.4)
Irinotecan + Cisplatin/Irinotecan + Carboplatin	46 (36.2)
Pemetrexed + Cisplatin/Pemetrexed + Carboplatin	36 (28.3)

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects, (a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

The lung cancer patients were also classified according to RECIST criteria into ECOG and KPS. According to ECOG the study subjects were classified as 86.9% (214) subjects with ECOG score 0-1, 12.6% subjects with ECOG score >2. For performance parameter KPS there were 63.0% subjects with KPS score between 100-80, 32.9% subjects with KPS score <70%. 162 of the study subjects were administered platinum based chemotherapy as first line treatment. 35.4% (45) of the study subjects were given docetaxel along with cisplatin/carboplatin. 36.2% (46) study patients were administered irinotecan along with cisplatin/carboplatin. 28.3% (36) subjects were given pemetrexed in combination with cisplatin/carboplatin.

5.12.2 Association of *XRCC4* variant with overall survival of lung cancer patients

Table 5.75 shows the overall survival and survival according to histological stratification of subjects with polymorphic *XRCC4*. Our data analysis suggested no affect on the overall survival for the genotypic combinations of polymorphic *XRCC4* *Intron 3* and lung cancer patients. The frequency distribution suggested there were 26.8% subjects with wild type genotype *II*. 46.3% subjects with heterozygous genotype (*ID*) and 26.8% subjects with mutant genotype (*DD*). The median survival time for double allelic variants (*DD*) was 6.2 months compared to 6.1 months for heterozygous subjects (*ID*) and 5.0 months for subjects with wild type genotype (*II*). No difference in the death rate among the genotypes was observed.

Further survival analysis of subjects with histological stratification was also evaluated in relation to *XRCC4* polymorphism as depicted in table 5.75. Subjects diagnosed with ADCC subtype and with insertion in the both alleles (*II*) had a higher MST compared to subjects with deletion in one allele (*ID*) (8.8 vs 6.3 months, log-rank p-value=0.33). Hazard rate by multivariate Cox-regression analysis did not show any association with survival. Similarly, SQCC subjects with single allelic deletion (*ID*) had a survival of 5.9 months compared to 4.3 months for wild type subjects (*II*) (Log rank p-value=0.51), whereas subjects with mutant (*DD*) genotype had a MST of 4.3 months (log rank p-value=0.81). Subjects diagnosed with SCLC subtype revealed the highest MST for mutant subjects (*DD*) (MST=8.2 months, log rank p-value=0.47) compared to heterozygous (*ID*) and wild type subjects(*II*) ((*ID*):MST=6.7, log rank p-value=0.66, (*II*):MST=4.1). Cox-regression analysis for subjects diagnosed with SCLC subtype showed a trend of high hazard rate. Subjects with heterozygous genotype (*ID*) had a 1.6 fold hazard rate (H.R.=1.58, 95%C.I.=0.61-4.12, p=0.35). Similarly, subjects with mutant genotype (*DD*) had a higher hazard rate 1.97 (H.R=1.97, 95%C.I.=0.58-5.21, p=0.27) as shown in table 5.75.

Table 5.75: *XRCC 4* polymorphic variant and its association with overall survival

Genes	Cases N=246 (n%)	Univariate analysis			Multivariate analysis	
		MST (months)	Log rank p-value	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	p-value
XRCC4 Intron 3 I/D						
II	66 (26.8)	5.0		1	1	
ID	114 (46.3)	6.1	0.4	0.87	0.97 (0.66-1.41)	0.88
DD	66 (26.8)	6.2	0.6	0.91	0.96 (0.78-1.88)	0.72
ID+DD	180 (73.1)	6.2	0.42	0.88	0.97 (0.69-3.18)	0.90
ADCC	N=92					
II	29 (31.5)	8.8		1	1	
ID	39 (42.3)	6.3	0.33	0.78	1.05 (0.56-1.94)	0.87
DD	24 (26.0)	8.0	0.80	1.07	0.98 (0.50-1.89)	0.96
ID+DD	63 (68.4)	7.13	0.44	1.19	1.02 (0.60-1.74)	0.92
SQCC	N= 94					
II	24 (25.5)	4.30		1	1	
ID	44 (46.8)	5.93	0.51	1.21	0.79 (0.41-1.53)	0.49
DD	26 (27.6)	5.43	0.81	1.08	-	-
ID+DD	70 (74.4)	5.76	0.57	1.17	0.80 (0.43-1.48)	0.48
SCLC	N=59					
II	13 (22.0)	4.10		1	1	
ID	30 (50.8)	6.73	0.66	1.16	1.58 (0.61-4.12)	0.35
DD	16 (27.1)	8.27	0.47	1.31	1.97 (0.58-6.62)	0.27
ID+DD	46 (77.9)	6.73	0.61	1.18	2.11 (0.85-5.21)	0.10

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.12.3 Association of *XRCC4* polymorphism and overall survival according to chemotherapy regimens given to lung cancer patients

Second line chemotherapy was given to lung cancer subjects as a primary treatment. Total 127 subjects were administered chemotherapy where out of them 35.4% (45) were administered cisplatin/carboplatin along with docetaxel. 36.2% (46) were given cisplatin/carboplatin in combination with irinotecan and remaining 28.3% (36) were given cisplatin/carboplatin along with pemetrexed. Subjects carrying deletion of one allele (ID) and treated with cisplatin/carboplatin along with docetaxel were represented more compared to subjects carrying insertion at both alleles (II) or deletion at both alleles (DD) (46.6 vs 26.6 and 26.6%) as shown in table 5.76. Subjects with deletion at one allele had a better MST compared to wild type subjects (MST=5.1 vs 5.0, log rank p-value=0.9). Cox regression revealed a trend towards increase in hazards rate (H.R.=1.87, 95% C.I.=0.58-6.0, p=0.29). Subjects administered cisplatin/carboplatin in combination with pemetrexed and harboring deletion of one allele (ID) (50.0%) were represented more compared to patients carrying homozygous wild (II) (22.2%) or mutant genotype (27.7%). Subjects carrying deletion of one allele (ID) had a MST of 5.4 months with no association with death rate (H.R.=0.88, 95%C.I.=0.36-2.1) as shown in table 5.78. When hazards rate was adjusted for factors age, sex, smoking, histology, stage, KPS, ECOG hazard rate reduced further (H.R.=0.2, 95%C.I.=0.05-1.1, p=0.06) (figure 5.25b). Subjects carrying deletion of both alleles had a better survival of 21.0 months with a hazard rate of 0.5 fold (H.R.=0.5, 95%C.I.=0.19-2.6, p=0.28).

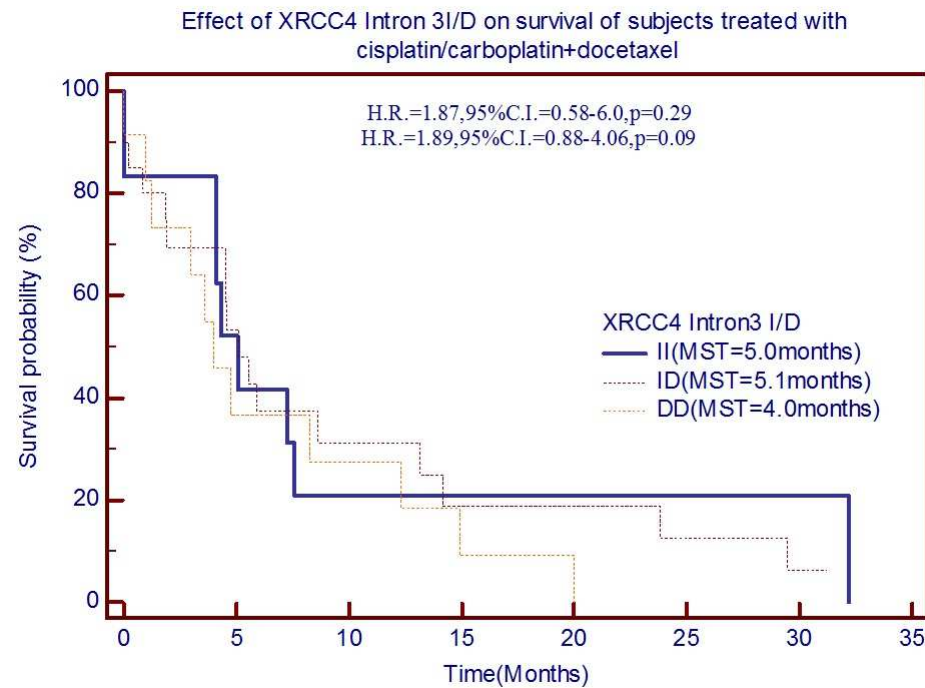
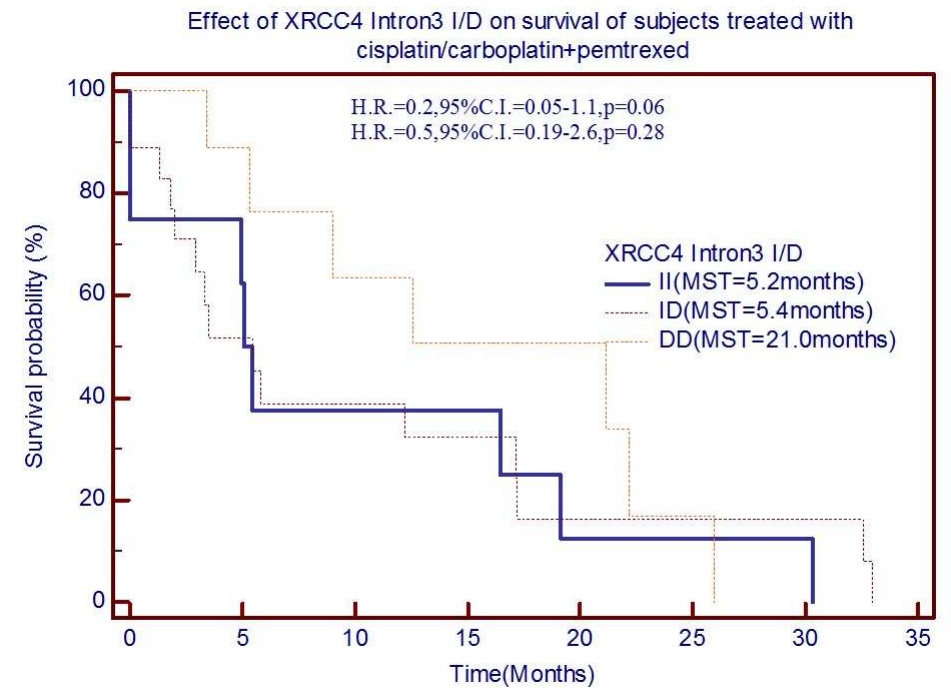
(a) *XRCC4* intron 3 I/D(b) *XRCC4* intron 3 I/D

Figure 5.25: Kaplan Meier Survival Curve for XRCC4 intron3 I/D variant in lung cancer subjects treated with (a) Cisplatin/Carboplatin +Docetaxel (II:reference) (b)Cisplatin/Carboplatin +Pemetrexed (II:reference)

Table 5.76: Association of *XRCC4* SNP with chemotherapy regimen (Cisplatin / Carboplatin+Docetaxel) and overall Survival

Genes	Cisplatin/Carboplatin+Docetaxel					
XRCC4 INTRON 3(I/D)	Cases n(%) N=45	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
II	12 (26.6)	5.0		1	1	
ID	21 (46.6)	5.1	0.9	0.98 (0.43-2.2)	1.87 (0.58-6.0)	0.29
DD	12 (26.6)	4.0	0.6	1.23 (0.51-2.97)	1.89 (0.88-4.06)	0.09
ID+DD	33 (73.3)	4.7	0.84	1.07 (0.5-2.2)	1.55 (0.56-4.28)	0.39

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.77: Association of *XRCC4* SNP with chemotherapy regimens (Cisplatin / Carboplatin+Irinotecan) overall Survival

Genes	Cisplatin/Carboplatin+Irinotecan					
XRCC4 Intron 3 (I/D)	Cases n(%) N=46	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
II	14 (30.4)	12.0			1	
ID	19 (41.3)	9.0	0.87	1.06 (0.45-2.5)	1.16 (0.37-3.7)	0.79
DD	13 (28.2)	11.5	0.95	0.9 (0.40-2.31)	0.89 (0.48-1.64)	0.71
ID+DD	32 (69.5)	9.8	0.95	1.01 (0.47-2.17)	0.98 (0.41-2.30)	0.96

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

Table 5.78: Association of *XRCC4* SNP with chemotherapy regimens (Cisplatin / Carboplatin+Pemetrexed) and overall Survival

Genes	Cisplatin/Carboplatin+Pemetrexed					
XRCC4 Intron 3 (I/D)	Cases n(%) N=36	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
II	8 (22.2)	5.2				
ID	18 (50.0)	5.4	0.76	0.88 (0.36-2.1)	0.2 (0.05-1.1)	0.06
DD	10 (27.7)	21.0	0.4	0.68 (0.24-1.90)	0.5 (0.19-2.6)	0.28
ID+DD	28 (77.7)	5.8	0.49	0.76 (0.31-1.89)	0.23 (0.06-0.88)	0.02

(a)-Unadjusted Hazards ratio for Kaplan meier analysis, (b)-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.12.4 Association of *XRCC4* polymorphism and its correlation with objective response

The objective response was evaluated for the polymorphic variant of *XRCC4* located on intron 3. The percentage of subjects having deletion of single allele (ID) was 55.8% in good responders compared to 38.9% in poor responders. There were 22.0% subjects with wild type (II) genotype in good responders compared to 33.8% subjects in poor responders. Logistic regression analysis revealed a 7 fold OR for heterozygous subjects giving a good response as compared to patients having wild type genotype (II) (OR=7.34, 95%CI=1.31-41.1, **p=0.02**).

Table 5.79: Effect of *XRCC4* genetic variants with treatment response

Genotype	Good response CR+PR n(%) N = 68	Poor response SD+PD n (%) N = 59	AOR (95% CI) ^a	p ^b
XRCC4 ID				
II	15 (22.0)	20 (33.8)	1	
ID	38 (55.8)	23 (38.9)	7.34 (1.31-41.1)	0.02
DD	15 (22.0)	16 (27.1)	1.51 (0.74-3.05)	0.24
ID+DD	53 (77.9)	39 (66.1)	2.24 (0.63-7.95)	0.21

(a)-Adjusted Odds ratio, (b)- 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, histology, stage, KPS, ECOG. Abbreviations: CR-Complete reponse, PR-Partial reponse, SD-Stable disease, PD-Progressive disease, AOR-Adjusted odds ratio, C.I.-Confidence interval

5.13 Survival analysis of DSBRR genes *XRCC6* and *XRCC7*

5.13.1 Clinical characteristics of subjects with polymorphic *XRCC6* and *XRCC7*

The survival analysis of lung cancer patients in relation to *XRCC6* 61 C>G and *XRCC7* 6721 G>T polymorphism was also evaluated as illustrated in table 5.80. The common sample size for the genes together was 241. The median age for the study group was 60±10.7 (58-61). The distribution based on gender suggested there were 88.7% (214)

males and 11.2% (27) females. The study further elucidated there were 81.7% (197) smokers and 18.2% (44) non-smokers. The classification based on histology suggested there were 34.4% (82) ADCC patients, 41.0% (99) SQCC patients and 23.6% (57) SCLC subjects and remaining 0.8% (2) were unclassified. The TNM staging classified the study into 53.5% (129) stage III and 45.2 % (109) stage IV subjects whereas remaining 1.2% (3) subjects were unclassified. According to the performance status ECOG there were 52.2% (126) subjects with ECOG score between 0-1, 29.8% (72) with ECOG 2, and 17.8% (43) subjects with ECOG 3-4. According to KPS status there were 60.1% (145) subjects with KPS score of 100-80, 31.9% (77) subjects with KPS score 70-60 and 7.8% (19) with KPS score of 50-40. There were 197 (81.7%) deaths or events recorded. The chemotherapeutic treatment was given to 162 study subjects where out of them there were 39.5% (64) subjects were given docetaxel + cisplatin/carboplatin. 26.5% (43) study subjects were treated with irinotecan + cisplatin/carboplatin, 33.9% (55) subjects were given pemetrexed + cisplatin/carboplatin.

Table 5.80: Patients Characteristics and clinical features

Patients characteristics	No.of Patients, n(%) N=241
Age(Years)	
Median±SD Range	60±10.7 (58-61)
Gender	
Male	214 (88.7)
Female	27 (11.2)
Smoking status	
Smokers	197 (81.7)
Non-smokers	44 (18.2)
Histology	
ADCC	83 (34.4)
SQCC	99 (41.0)
SCLC	57 (23.6)
Others	2 (0.8)
TNM staging	
I	2 (0.8)
II	8 (3.3)
III	119 (49.3)
IV	109 (45.2)
Unclassified	3 (1.2)
ECOG performance status	
0-1	126 (46.7)
2	72 (29.8)
3-4	43 (17.8)
KPS performance status	

to be cont'd on next page

Table 5.80: Patients Characteristics and clinical features of XRCC6 and XRCC7

100-80	145 (60.1)
70-60	77 (31.9)
50-40	19 (7.8)
Events, Death	197 (81.7%)
Chemotherapy Regimen	N =162
Docetaxel + Cisplatin/Docetaxel + Carboplatin	64 (39.5)
Irinotecan + Cisplatin/Irinotecan + Carboplatin	43 (26.5)
Pemetrexed + Cisplatin/Pemetrexed + Carboplatin	55 (33.9)

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects, (a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age. All p-values are two sided p<0.05 was considered statistically significant

5.13.2 Association of polymorphic *XRCC6*, *XRCC7* with overall survival of lung cancer patients

The survival analysis for *XRCC6* 61C>G and *XRCC7* 6721G>T revealed that none of the two SNPs showed any significant difference in the median survival time for any genotype. However, the median survival time for subjects carrying double allelic variant (GG) genotype was equal to that of patients with heterozygous genotype (CG) (GG=5.6 vs CG=5.5). On the contrary patients carrying the wild type genotype (CC) had a MST of 8.6 as indicated in figure 5.26a. When Cox regression was applied *XRCC6* 61C>G did not show any effect on death rate. In case of *XRCC7* 6721G>T subjects with mutant genotype (TT) showed a MST of 6.0 months while single allelic carriers (GT) showed a MST of 7.1 months (figure 5.26b). Cox-regression analysis for single allelic variants (GT) in case of *XRCC7* 6721G>T depicted a hazard ratio of 0.68 (H.R=0.68, 95%C.I.=0.46-1.00, **p=0.04**) which was significant as shown in table 5.81.

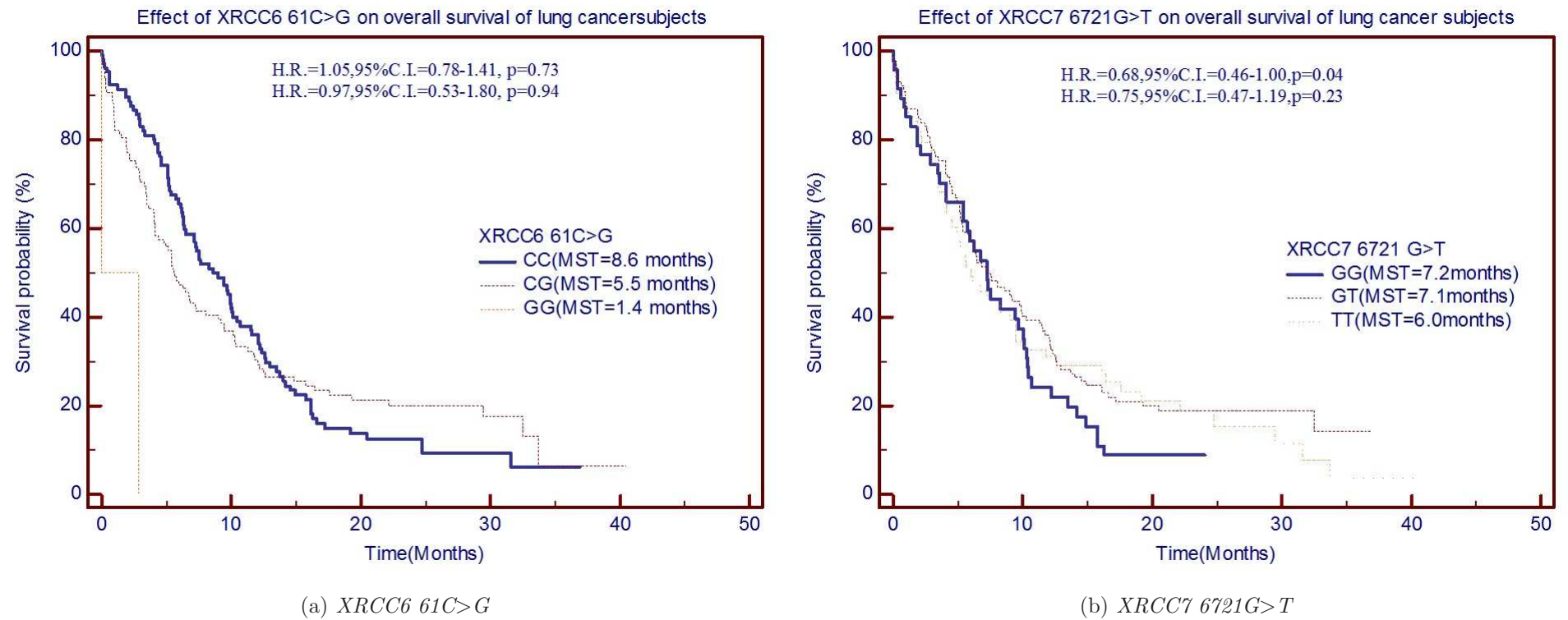


Figure 5.26: Kaplan Meier Overall Survival Curve for (a) *XRCC6 61C>G* (CC:reference) (b) *XRCC7 6721G>T* (GG:reference).
H.R.=Hazard rate, MST=Median survival time

Table 5.81: *XRCC6 61C>G* and *XRCC7 6721G>T* polymorphic variants and their association with overall survival

Genes	Cases N=241 (n%)	Death N=197 (n%)	Alive N=44 n(%)	Univariate analysis			Multivariate analysis	
				MST (months)	Log rank p-value	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	p-value
XRCC6 61C>G								
CC	105 (43.5)	90 (45.6)	15 (34.0)	8.6		1	1	
CG	120 (49.7)	95 (48.2)	25 (56.8)	5.5	0.7	1.05 (0.79-1.40)	1.05 (0.78-1.41)	0.73
GG	16 (6.6)	12 (6.0)	4 (9.0)	5.6	0.78	0.94 (0.70-1.24)	0.97 (0.53-1.80)	0.94
CG+GG	136 (56.4)	107 (54)	29 (65.9)	5.5	0.67	0.91 (0.49-1.71)	0.98 (0.72-1.34)	0.92
XRCC7 6721G>T								
GG	47 (19.5)	42 (21.3)	5 (11.3)	7.2		1	1	
GT	131 (54.3)	102 (51.7)	29 (65.9)	7.1	0.17	0.78 (0.53-1.14)	0.68 (0.46-1.00)	0.04
TT	63 (26.1)	53 (26.9)	10 (22.7)	6.0	0.38	1.27 (0.87-1.87)	0.75 (0.47-1.19)	0.23
GT+TT	194 (80.4)	155 (78.6)	39 (88.6)	6.8	0.20	1.18 (0.78-1.79)	0.71 (0.50-1.03)	0.07

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

5.13.3 Cox-regression analysis of the factors associated with overall survival of *XRCC7*

Table 5.82: Results of cox-regression multivariate analysis for subjects with polymorphic *XRCC7 6721G>T*

Variables	B	SE	Wald	P-value	Exp(b)	95%C.I.
<i>XRCC7 6721G>T</i>	-0.380	0.195	3.771	0.052	0.683	0.466 - 1.001
AGE	0.0122	0.008	1.913	0.166	1.012	0.995 - 1.029
SEX	-0.444	0.324	1.872	0.171	0.641	0.340 - 1.207
Smoking	0.032	0.313	0.011	0.916	1.033	0.561 - 1.902
Histology	0.309	0.138	4.981	0.025	1.362	1.039 to 1.784
Stage	0.285	0.144	3.893	0.048	1.330	1.003 - 1.765
ECOG	-0.041	0.193	0.045	0.831	0.959	0.657 - 1.399
KPS	-0.207	0.219	0.892	0.344	0.812	0.5297 - 1.2468

B-Regression coefficient; SE-Standard error; Wald-Wald value;HR-Hazard ratio; 95% C.I.-confidence intervals

The multivariate regression analysis showed that age, sex, smoking, ECOG, KPS were not associated with the prognosis for heterozygous subjects in case of polymorphic *XRCC7 6721G>T*. Histology and stage for subjects carrying heterozygous genotype in case of *XRCC7 6721G>T* were independent prognostic factor for the 3-year survival rate.

5.13.4 Association of polymorphic *XRCC6*, *XRCC7* and survival according to histological subtypes of lung cancer

Overall survival was further assessed in lung cancer patients according to histological stratification. For polymorphic variant of *XRCC6* subjects diagnosed with ADCC subtype there were 28.9% subjects carrying the wild type (*CC*) genotype, 49.3% of patients were heterozygous (*CG*) subjects and 21.6% of the patients were having the mutant genotype (*GG*). Subjects carrying mutant genotype (*GG*) and with ADCC subtype showed a MST of 9.4 months as compared to MST of 7.6 and 6.3 months for heterozygous (*CG*) and wild type (*GG*) genotype respectively as shown in table 5.83. Cox-regression analysis suggested no effect on lung cancer as hazard rate was nearly 1 for every genotypic combination for both *XRCC6 61C>G* and *XRCC7 6721G>T* polymorphic sites.

It was further observed for polymorphic variant *XRCC6 61C>G* that subjects diagnosed

with SCLC histologic subtype and carrying mutant genotype GG were the least represented (3.5%) as compared to CG (40.3%) and CC (56.1%). Survival analysis suggested subjects carrying both the alleles of the mutant genotype (GG) for $XRCC6\ 61C>G$ had a poor survival rate as they survived for less number of months as compared to patients carrying the wild type genotype (CC) (MST=9.8 vs 1.4, log rank p-value=**0.0017**) (figure 5.27a) and there was a significant difference between the number of months survived. Further when data was adjusted by cox regression model for factors such as age, sex, smoking, histology, stage, KPS, ECOG a very high death rate was observed of 23 fold (H.R=23.0, 95%C.I=3.30-172.95, **$p=0.0017$**) for mutant genotype (GG) suggesting a poor prognosis. Similarly, patients with single allelic variant genotype (CG) showed a high death rate (H.R=1.89, 95%C.I=0.96-3.73, $p=0.06$) as shown in table 5.85.

In case of $XRCC7\ 6721G>T$ frequency distribution for SCLC subtype suggested there were 7.0% subjects carrying wild type (GG) genotype, 56.1% subjects were heterozygous (GT), and 36.8% subjects were with mutant (TT) genotype. The median survival time for subjects with mutant genotype (TT) was higher compared to subjects with wild genotype (GG) (MST= 7.3 vs 5.3, log rank p-value=0.21). Therefore, univariate analysis did not suggest any significant difference in the median survival time figure 5.27b. The death rate for mutant subjects indicated no a trend of no association with survival (H.R=0.74, 95%C.I.=0.18-2.34, $p=0.07$). Cox-regression analysis further reduced a trend of small death rate for combined variant genotype ($GT+TT$) (H.R=0.49, 95%C.I=0.16-1.43, $p=0.19$) as shown in table 5.85.

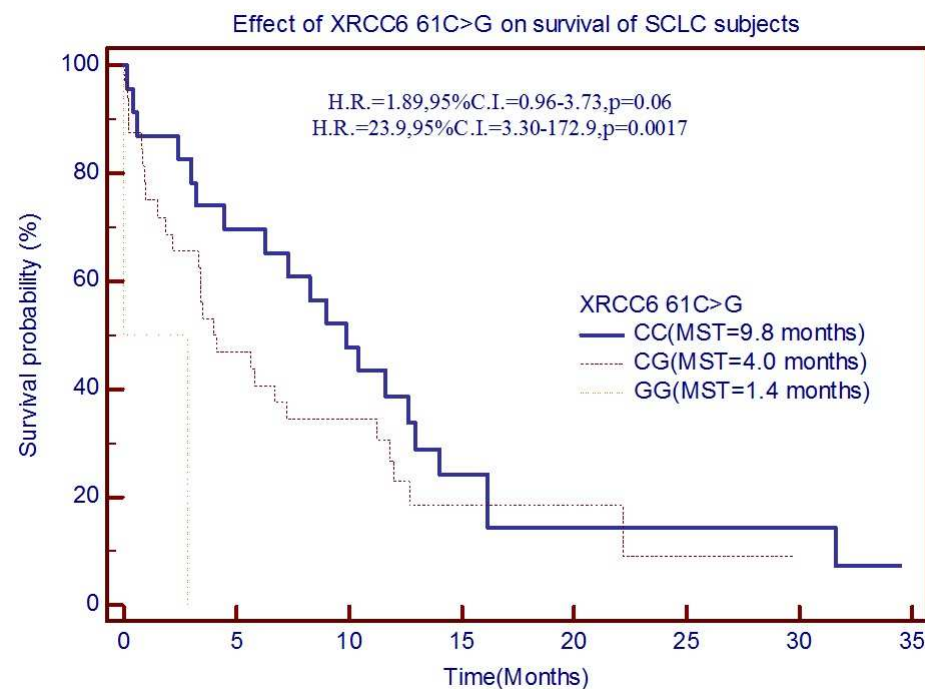
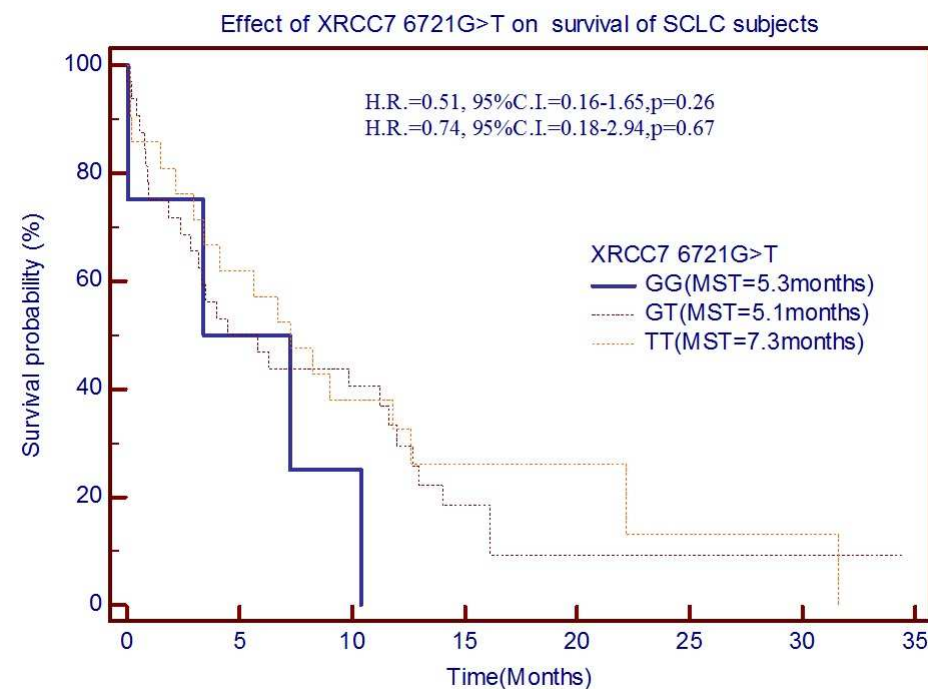
(a) *XRCC6 61C>G*(b) *XRCC7 6721 G>T*

Figure 5.27: Kaplan Meier Overall Survival Curve for XRCC6 and XRCC7 variants in lung cancer subjects stratified according to their histology (a) *XRCC6 61C>G* (CC:reference) in SCLC subjects (b) *XRCC7 6721 G>T* (GG:reference) in SCLC subjects

Table 5.83: Association of polymorphic *XRCC6*, *XRCC7* with Overall Survival according to Tumor histology ADCC

Genes	ADCC					
XRCC6 61 C>G	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	24 (28.9)	6.3			1	
CG	41 (49.3)	7.6	0.2	1.32 (0.78-2.2)	1.28 (0.73-2.25)	0.38
GG	18 (21.6)	9.4	0.75	1.16	1.00 (0.37-2.73)	0.98
CG+GG	59 (71.08)	8.3	0.24	1.29 (0.79-2.12)	1.24 (0.72-2.1)	0.43
XRCC7 6721 G>T	Cases n(%) N=83	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	24 (75.9)	7.5			1	
CT	43 (22.8)	7.6	0.53	0.84 (0.47-1.48)	0.80 (0.44-1.47)	0.48
TT	16 (1.2)	6.0	0.71	0.88 (0.44-1.77)	0.80 (0.38-1.67)	0.56
CT+TT	59 ()	7.1	0.53	0.84 (0.49-1.46)	0.87 (0.49-1.54)	0.64

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.84: Association of polymorphic *XRCC6*, *XRCC7* with Overall Survival according to Tumor histology SQCC

Genes	SQCC					
XRCC6 61 C>G	Cases n(%) N=99	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	39 (39.3)	8.6			1	
CG	53 (53.5)	6.2	0.17	0.73 (0.46-1.16)	0.71 (0.44-1.13)	0.16
GG	7 (7.0)	8.0	0.48	0.72 (0.31-1.65)	0.60 (0.28-1.72)	0.34
CG+GG	60 (60.6)	6.2	0.17	0.74 (0.47-1.16)	0.72 (0.45-1.43)	0.15
XRCC7 6721 G>T	Cases n(%) N=99	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	19 (19.1)	7.2			1	
CT	54 (54.5)	8.0	0.16	0.68 (0.37-1.24)	0.58 (0.31-1.07)	0.05
TT	26 (26.2)	5.2	0.63	0.86 (0.46-1.62)	0.81 (0.41-1.62)	0.56
CT+TT	80 (80.8)	6.8	0.23	0.74 (0.41-1.31)	0.66 (0.38-1.14)	0.14

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.85: Association of polymorphic *XRCC6*, *XRCC7* with Overall Survival according to Tumor histology SCLC

Genes	SCLC					
	Cases n(%) N=57	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^b	H.R. ^b	p-value
CC	23 (40.3)	9.8			1	
CG	32 (56.1)	4.0	0.28	1.36 (0.76-2.43)	1.89 (0.96-3.73)	0.06
GG	2 (3.5)	1.4	0.0017	7.03 (0.20-24.3)	23.92 (3.30-172.95)	0.002
CG+GG	34 (59.6)	3.5	0.20	1.43 (0.81-2.52)	2.03 (1.04-3.97)	0.03
XRCC7 6721 G>T	Cases n(%) N=57	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	4 (7.0)	5.3			1	
CT	32 (56.1)	5.1	0.28	0.57 (0.15-2.12)	0.51 (0.16-1.65)	0.26
TT	21 (36.8)	7.3	0.21	0.51 (0.13-2.02)	0.74 (0.18-2.94)	0.67
CT+TT	53 (92.9)	6.3	0.23	0.54 (0.14-2.07)	0.49 (0.16-1.43)	0.19

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

5.13.5 Role of polymorphic *XRCC6* and *XRCC7* gene and overall survival according to chemotherapy regimens administered

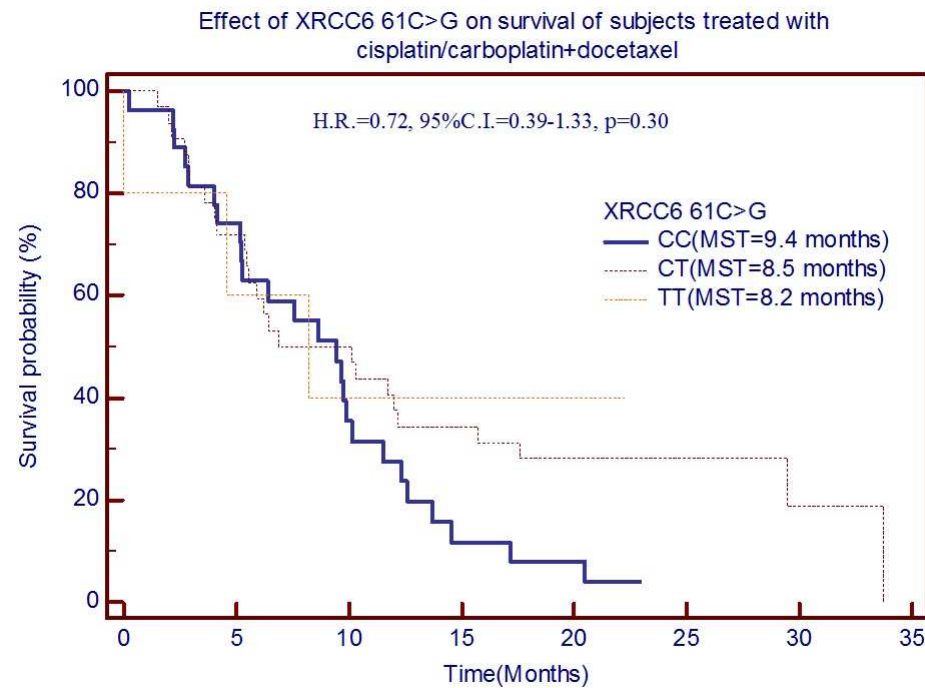
Study data was further assessed according to the three chemotherapeutic regimens given to the study subject. The study group was classified according to three regimens given to the study subjects. First group was administered cisplatin/carboplatin along with docetaxel. It was observed that for *XRCC6 61C>G*, percentage of subjects carrying combined variant (*CG+GG*) was 57.8%. The median survival time was 8.2 months for the group with p-value of nearly approaching significance (0.08). Cox-regression for the same polymorphic gene did not show any influence on death rate as shown in table 5.86 (figure 5.28a).

In case of patients receiving cisplatin/carboplatin along with docetaxel and carrying single allelic variant genotype it was observed that the frequency distribution for *XRCC7* showed that 17.1% lung cancer patients had wild type genotype (*GG*), 60.9% subjects

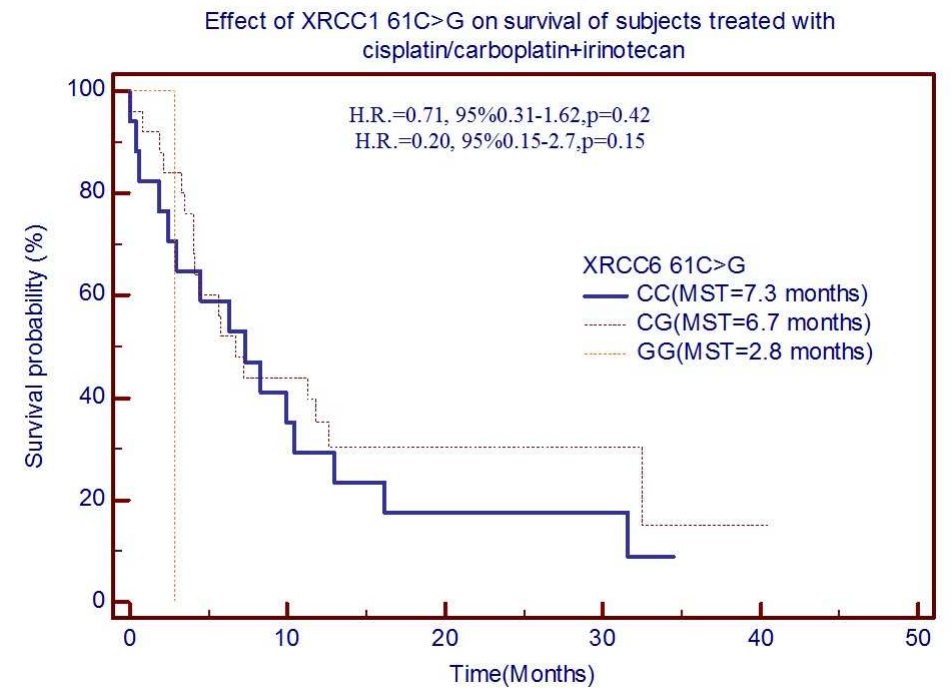
were carriers for heterozygous genotype (*GT*), and remaining 21.8% subjects had mutant (*TT*) genotype. Subjects with single allelic variants (*GT*) had a MST of 9.8 months as compared to MST of 5.4 months in case of subjects with wild type genotype (*GG*) (**log rank $p=0.03$**) (figure 5.29a). Cox-regression analysis revealed a better prognosis with very small hazard rate (H.R.=0.40, 95%C.I.=0.15-1.03, $p=0.05$). Similarly, when heterozygotes (*GT*) and mutant genotype (*TT*) were combined as a single genotype a MST of 9.7 months was observed (**$p=0.04$**). Sequentially, when regression was applied it was observed that heterozygous genotype (*GT*) had a lower death rate (H.R.=0.44, 95%C.I.=0.19-0.99, $p=0.05$).

For subjects administered cisplatin/ carboplatin along with irinotecan, it was reported that for *XRCC6 61C>G* the distribution frequency revealed there were 39.5% subject harboring homozygous wild genotype (*CC*), 58.1% were of heterozygous genotype (*CG*) and 2.3% were of mutant genotype *GG*. The median survival time for mutant genotype (*GG*) (MST=2.8), on comparison with wild type (*GG*) subjects (MST=7.3), it was not significant (log rank p -value=0.31) as shown in figure 5.28b. Subjects with polymorphic *XRCC7 6721G>T* and administered cisplatin/ carboplatin in combination with irinotecan suggested there was a higher representation of single allelic carriers (*GT*) 58.1% as compared to wild (*GG*) and mutant subjects (*TT*) subjects (6.9 vs 34.8%). Statistical analysis suggested the median survival time for wild type (*GG*) genotype which was 7.2 months compared to 5.8 months for single allelic variants (*GT*) and 7.3 months for mutant (*TT*) subjects as indicated in table 5.87 (figure 5.29b). Cox-regression analysis suggested a protective effect for double allelic variants (*TT*) (H.R.=0.34, 95%C.I.=0.05-2.35, $p=0.28$).

For both polymorphic variants *XRCC6 61C>G* and *XRCC7 6721G>T* the survival analysis did not reveal any correlation with survival for subjects administered cisplatin/ carboplatin + pemetrexed.

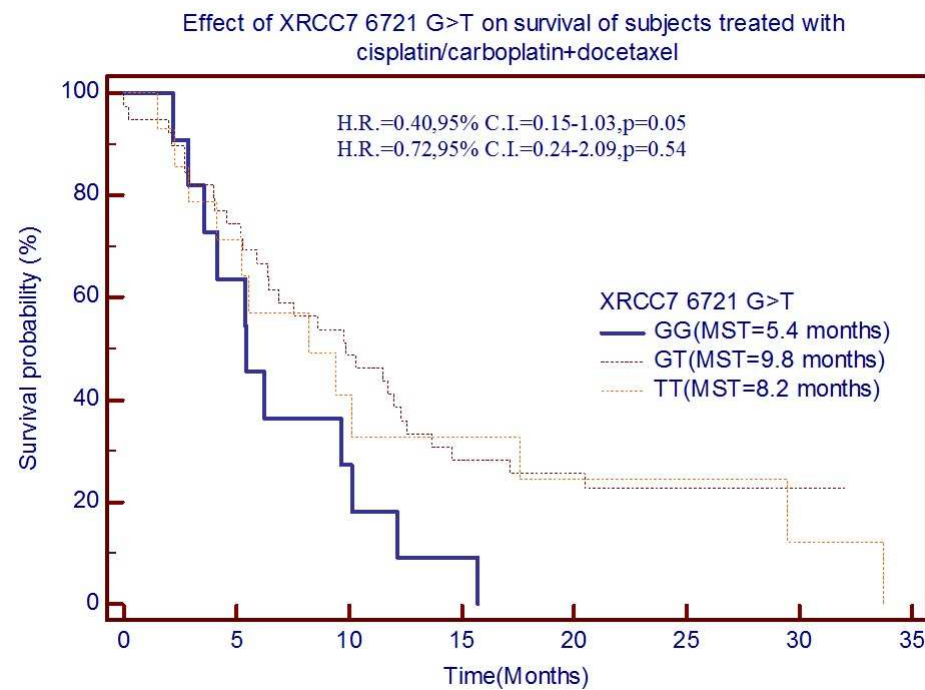


(a) Cisplatin/Carboplatin+Docetaxel

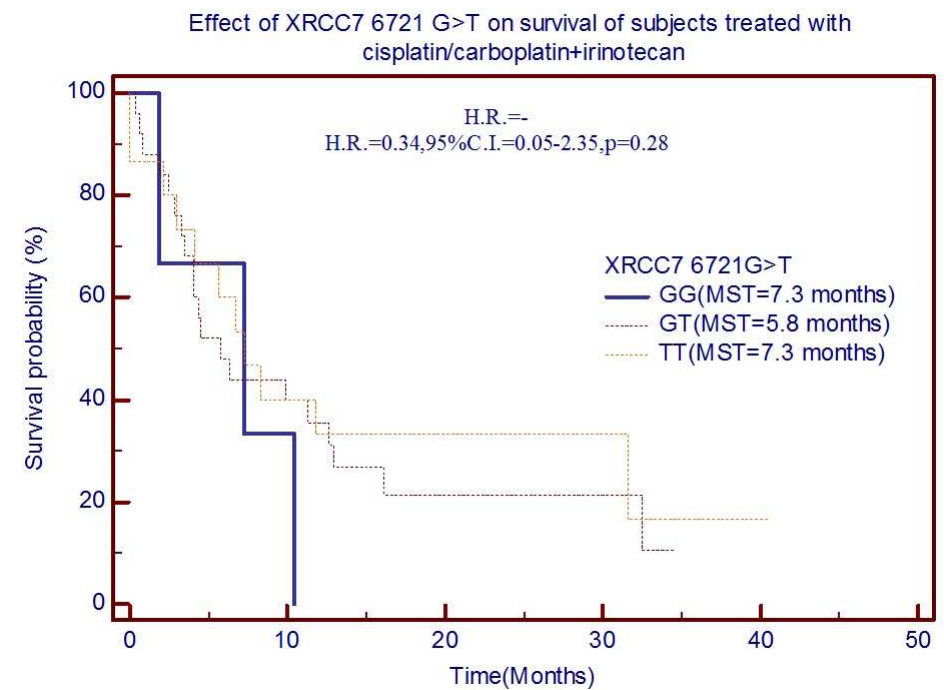


(b) Cisplatin/Carboplatin+Irinotecan

Figure 5.28: Kaplan Meier overall survival Curve for *XRCC6* 61C>G subjects treated with (a) Cisplatin / Carboplatin + Docetaxel (b) Cisplatin / Carboplatin + Irinotecan



(a) Cisplatin/Carboplatin+Docetaxel



(b) Cisplatin / Carboplatin+Irinotecan

Figure 5.29: Kaplan Meier overall survival Curve for *XRCC76721 G>T* subjects treated with (a) Cisplatin / Carboplatin + Docetaxel (b) Cisplatin / Carboplatin +I rinotecan

Table 5.86: Polymorphic *XRCC6* and *XRCC7* and survival for subjects administered Cisplatin/Carboplatin+Docetaxel

Genes	Cisplatin/Carboplatin+Docetaxel					
XRCC6 61 C>G	Cases n(%) N=64	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	27 (42.1)	8.6		1		
CG	32 (50.0)	8.5	0.09	0.63 (0.35-1.12)	0.72 (0.39-1.33)	0.30
GG	5 (7.8)	8.2	0.25	0.51 (0.19-1.31)	-	-
CG+GG	37 (57.8)	8.2	0.08	0.62 (0.35-1.11)	0.66 (0.36-1.19)	0.17
XRCC7 6721 G>T	Cases n(%) N=64	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	11 (17.1)	5.4		1		
GT	39 (60.9)	9.8	0.03	0.49(0.21-1.16)	0.40 (0.15-1.03)	0.05
TT	14 (21.8)	8.2	0.17	0.59(0.25-1.39)	0.72 (0.24-2.09)	0.54
GT+TT	53 (82.8)	9.7	0.04	0.52(0.22-1.19)	0.44 (0.19-0.99)	0.05

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.87: Polymorphic *XRCC6* and *XRCC7* and survival for subjects administered Cisplatin/Carboplatin+Irinotecan

Genes	Cisplatin/Carboplatin+Irinotecan					
XRCC6 61 C>G	Cases n(%) N=43	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	17 (39.5)	7.3	1	1		
CG	25 (58.1)	6.7	0.48	0.78 (0.43-1.67)	0.71 (0.31-1.62)	0.42
GG	1 (2.3)	2.8	0.31	2.0 (0.11-63.20)	0.20 (0.15-2.7)	0.15
CG+GG	26 (60.4)	6.7	0.54	0.81(0.40-1.62)	0.75 (0.33-1.72)	0.51
XRCC7 6721 G>T	Cases n(%) N=43	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	3 (6.9)	7.2		1		
GT	25 (58.1)	5.8	0.4	0.62 (0.14-2.69)	-	-
TT	15 (34.8)	7.3	0.37	0.57 (0.12-2.64)	0.34 (0.05-2.35)	0.28
GT+TT	40 (93.0)	6.3	0.39	0.60 (0.13-2.6)	0.76 (0.19-3.0)	0.70

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

Table 5.88: Polymorphic *XRCC6* and *XRCC7* and survival for subjects administered Cisplatin/Carboplatin+Pemetrexed

Genes	Cisplatin/Carboplatin+Pemetrexed					
XRCC6 61 C>G	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
CC	26 (47.2)	10.0		1		
CG	23 (41.8)	6.7	0.18	1.53 (0.78-3.03)	1.25 (0.51-3.05)	0.62
GG	6 (10.9)	7.5	0.88	1.08 (0.35-3.26)	-	-
CG+GG	29 (52.7)	6.7	0.26	1.42 (0.51-3.05)	0.93 (0.42-2.08)	0.87
XRCC7 6721 G>T	Cases n(%) N=55	Univariate analysis			Multivariate analysis	
		MST	p-value	H.R. ^a	H.R. ^b	p-value
GG	12 (21.8)	7.5		1		
GT	30 (54.5)	9.1	0.83	1.0 (0.49-2.45)	0.63 (0.29-1.94)	0.3
TT	13 (23.6)	6.1	0.50	1.3 (0.55-3.3)	1.28 (0.43-3.77)	0.65
GT+TT	43 (78.1)	7.6	0.64	1.19 (0.52-2.49)	0.71 (0.31-1.62)	0.42

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG

5.13.6 Association of *XRCC6* 61C>G and *XRCC7* 6721G>T polymorphisms with treatment response for different chemotherapeutic regimens and also their correlation with objective response

The objective response of the lung cancer subjects towards chemotherapeutic treatment was evaluated in relation to the genotype of polymorphic *XRCC6* and *XRCC7* has been summarized in table 5.89. On the basis RECIST criteria, the patients were stratified as good responders which included patients showing complete response (CR) and partial response (PR) and bad responders comprising patients showing progressive disease and stable disease. The good and poor responders were evaluated for every genotypic category for both polymorphic variant. Logistic regression analysis was used to evaluate the association of genotype with objective response where adjusted odds ratio was obtained after adjusting for various confounders such as age, gender, smoking, ECOG, KPS, stage and histology and chemotherapeutic regimens. In case of *XRCC6* 61 C>G the frequency distribution stated that single allelic carriers (CG) were almost similar in frequency in both good and poor responders (49.3% vs 48.3%). However, subjects carrying mutant genotype (GG) were 8.2% in good responders compared to 4.8% in poor responders. The association analysis exhibited a trend where a two fold good response as compared to patients having wild type genotype (CC) (OR=2.2, 95%C.I.=0.41-11.9, p=0.35). In case

of *XRCC7 6721 G>T* there were 58.0% and 56.1% single allelic carriers (*GT*) in good and poor responders respectively. Whereas, mutant subjects had higher representation in good response compared to subject in poor response (30.1% vs 22.5%). Regression analysis for both single allelic and double allelic variants did not show any association with lung cancer risk as shown in table 5.89.

Table 5.89: Effect of Polymorphic *XRCC6* and *XRCC7* on treatment response

Genotype	Good response CR+PR n(%) N = 73	Poor response SD+PD n (%) N = 62	AOR (95%C.I.) ^a	p ^b
<i>XRCC6 61 C>G</i>				
CC	31 (42.4)	29 (46.7)	1	
CG	36 (49.3)	30 (48.3)	0.96 (0.45-2.07)	0.93
GG	6 (8.2)	3 (4.8)	2.2 (0.41-11.9)	0.35
CG+GG	42 (57.5)	33 (53.2)	1.0 (0.52-2.3)	0.8
<i>XRCC7 6721 G>T</i>				
GG	10 (13.6)	12 (19.3)	1	
GT	41 (56.1)	36 (58.0)	1.2 (0.41-3.5)	0.72
TT	22 (30.1)	14 (22.5)	1.17 (0.61-2.27)	0.62
GT+TT	63 (86.3)	50 (80.6)	0.50 (0.67-2.23)	0.50

(a)-Adjusted Odds ratio, (b)- 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, histology, stage, KPS, ECOG. Abbreviations: CR-Complete reponse, PR-Partial reponse, SD-Stable disease, PD-Progressive disease, AOR-Adjusted odds ratio, C.I.-Confidence interval

5.14 Relationship of multiple gene interaction analysis among (*OGG1*, *MUTYH*, *XRCC1*, *XRCC4*, *XRCC6*, *XRCC7*) with lung cancer risk

5.14.1 Demographic characteristics

The multiple gene interaction analysis was carried for the sample population to unravel the interactive potential of these polymorphic sites in modulating lung cancer predisposition and also detecting their synergistic role in predicting clinical outcomes. However, it was a constraint to obtain the genotypic data for each sample for all the polymorphic

sites. So, the size of the study population was reduced to 302 cases and 302 controls as for these subjects the genotyping data was available for all polymorphic sites.

Table 5.90: Demographic Characteristics among cases and controls

Variable	Cases N=302 (n%)	Controls N=302 (n%)	p ^a
Age (Years)			
Mean±SD Range	57.94±10.58	53.42±10.09	<0.0001
Gender			
Male	263 (87.0)	250 (82.7)	0.172
Female	39 (12.9)	52 (17.2)	
Smoking status			
Smokers	243 (80.4)	212 (70.1)	0.004
Non-smokers	59 (19.5)	90 (29.8)	
Pack Years			
Mean±SD	29.0±36.0	17.0±19.4	<0.0001
Histology			
ADCC	103 (34.1)		
SQCC	129 (42.7)		
SCLC	68 (22.5)		
Others	2 (0.6)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects

(a) p-values were derived from Pearson Chisquare test except age; Student t-test was used for age.

All p-values are two sided p<0.05 was considered statistically significant

Study population consists of 302 cases and 302 controls where in the mean age was 57.94±10.58 in cases and 53.42±10.09 in controls. The number of males were 263 (87.0%) among cases and 250 (82.7%) among controls. The number of smokers were higher in cases (80.4%) as compared to controls (70.1%) as smoking is a significant risk factor for lung cancer. The patients were further stratified on the basis of their histological subtype and it was observed that 34.1% of the patients were classified into ADCC category, 42.7% fall under SQCC subgroup and 22.5% fall into SCLC category.

5.14.2 Genotypic distribution of SNP-SNP combinations of polymorphisms between BER and DSBR genes and their association with lung cancer predisposition

Table 5.91 shows the possible SNP-SNP combinations between the polymorphic sites of different genes namely *OGG1*, *MUTYH*, *XRCC1*, *XRCC4*, *XRCC6* and *XRCC7*. Logistic

regression was used to carry out the intergenic association analysis towards lung cancer. The genotypic frequencies, odd ratios, 95%C.I., and p-values are mentioned in the table 5.91. FDR adjusted p-values are also evaluated for the genotypic combinations and mentioned in table 5.91. The genotypic combination of *OGG1 Ser³²⁶Cys* and *XRCC4 intron 3* had a very strong association towards risk for lung cancer in the study population. The frequency of subjects carrying the heterozygous genotype (*CG+ID*) for the combination of both the snps was higher among cases as compared to control (87.1 vs 46.8%) group ($\chi^2=46.5$, $p\leq 0.0001$, $df=2$). Taking the wild genotype (*CC+II*) for both SNPs as a reference it was observed that single allelic carriers for both polymorphic sites were at very high risk towards lung cancer (OR=3.7, 95%C.I.=1.57-9.0, **$p=0.0030$** , $FDR_p=0.009$). On the same lines combined variants (*CG+ID*)+*GG+DD*) showed a high risk towards lung cancer when compared with the wild type genotype (OR=3.34, 95%C.I.=1.40-7.69, **$p=0.0063$** , $FDR_p=0.02$) as shown in table 5.91. The double combination of *OGG1 Ser³²⁶Cys* and *XRCC7 6721 G>T* exhibited an association towards lung cancer in North Indian population. Study subjects harboring single allele of variant genotype for both SNPs (*CG+GT*) represented a higher frequency in cases (88.8%) as compared to controls (68.9%) and subsequently a 2.5 fold risk was observed when compared with the wild type subjects (*CC+TT*) for the double combination (OR=2.47, 95%C.I.=1.0-6.05, **$p=0.047$** , $FDR_p=0.09$).

Further, for the next genotypic combination of *MUTYH Gln³²⁴His* and *XRCC4 intron 3* an association towards lung cancer was observed. It was observed that subjects carrying single alleles for both the SNPs (*GC+ID*) had a higher representation as compared to controls (63.8 vs 38.0%). When compared with wild type genotype for both the SNPs (*GG+II*) a risk of 3 fold was observed with statistical significance (OR=3.05, 95%C.I.=1.69-5.5, **$p=0.0002$** $FDR_p=0.002$). Adding to this, subjects with double allelic variant genotype (*CC+DD*) for the combination also showed an almost similar 3 fold risk (OR=2.94, 95%C.I.=1.63-5.30, **$p=0.0003$** , $FDR_p=0.002$) as shown in table 5.91.

The next double combination of *MUTYH Gln³²⁴His* and *XRCC6 61 C>G* showed a higher representation of heterozygous (*GC+CG*) genotype in cases (54.4%) as compared to controls (31.9%) ($\chi^2=14.8$, **$p=0.0006$** , $df=2$). Sequentially when compared to wild type (*GG+CC*) subjects a high risk towards lung cancer was observed (OR=2.41, 95%C.I.=1.39-4.18, **$p=0.0029$** , $FDR_p=0.009$). For the genotypic combination of *MUTYH Gln³²⁴His* and *XRCC7 6721 G>T* a higher percentage of subjects with heterozygous genotype (*GC+GT*) was reported among cases as compared to controls (67.8 vs 46.1%) ($\chi^2=12.7$, **$p=0.0017$** , $df=2$). Association analysis for the combination revealed that

carriers of single susceptible allele were at 2.3 fold of acquiring lung cancer (OR=2.32, 95%C.I.=1.26-4.27, $p=0.0068$, FDR_p=0.019) as detailed in table 5.91. However, this risk increased further for the double allelic variant genotype (GC+CC)+(GT+TT) (OR=2.45, 95%C.I.=1.36-4.39, $p=0.0026$, FDR_p=0.010).

The next double combination of *XRCC1 Arg³⁹⁹Gln* and *OGG1 Ser³²⁶Cys* showed diseased group had (76.7%) heterozygous subjects (GA+CG) while there were (46.8%) subjects under healthy control group. Henceforth, the heterozygous (GA+CG) genotype conferred a 2.5 fold increased risk in subjects compared to wild subjects (GG+CC) (OR=2.56, 95%C.I.=1.21-5.4, $p=0.013$, FDR_p=0.03) ($\chi^2=18.5$, $p=0.0001$, df=2). The next double combination of *XRCC1 Arg³⁹⁹Gln* and *MUTYH Gln³²⁴His* also exhibited a pronounced effect on lung cancer risk. Cases reported 30.8 % subjects with heterozygous genotype (GA+GC) and there were 17.5% subjects with heterozygous genotype (GA+GC) under controls. Risk analysis estimated an increased risk of 2 folds in patients having single allelic variant genotype (GA+GC) (OR=2.0, 95%C.I.=1.11-3.7, $p=0.02$, FDR_p=0.04) as shown in table 5.91. A higher frequency was observed for heterozygous genotype in cases (37.2%) for the combination of *XRCC1 399* and *XRCC4*, compared to controls (30.6 %), however the difference was not significant. Thereby, only a visible trend of high risk was observed for the combination (OR=1.64, 95%C.I.=0.83-3.2, p=0.15).

The genotypic combination of *XRCC1 Arg¹⁹⁴Trp* and *OGG1 Ser³²⁶Cys* also was found to modulate the lung cancer susceptibility of the study subjects. Subjects carrying heterozygous genotype (CT+CG) for both SNPs were reported to be higher in cases (67.8%) as compared to the controls (37.9%). The association analysis showed that these subjects were at 3 fold higher risk compared to those patients carrying the wild type genotype (CC+CC) (OR=3.07, 95%C.I.=1.57-5.9, $p=0.0010$, FDR_p=0.004). Similarly, heterozygous (CT+CG) subjects for the combination of *XRCC1 Arg¹⁹⁴Trp* and *MUTYH Gln³²⁴His* were higher represented in cases compared to controls (82.4 vs 75.7%) ($\chi^2=11.8$, $p=0.0006$, df=2). An odds ratio of at 3 fold was observed indicating a high risk (OR=3.13, 95%C.I.=1.61-6.08, $p=0.0007$, FDR_p=0.004) of acquiring lung cancer, when compared to the wild type genotype (CC+CC) as table 5.91 shows. In case of genotypic combination of *XRCC1 Arg¹⁹⁴Trp* and *XRCC4 intron 3* it was observed that the frequency of subjects carrying the heterozygous genotype was observed to be 34.4% in cases and 25.4% in controls. The difference in the frequency was not significant though a visible trend of high risk was observed (OR=1.45, 95%C.I.=0.74-2.82, p=0.27).

The next polymorphic variant of *XRCC1* gene *i.e.* *XRCC1 Pro²⁰⁶Pro* was evaluated along with *OGG1 Ser³²⁶Cys* and it was observed, that the difference in the frequency of subjects

carrying heterozygous genotype for cases and controls was significant (89.2% vs 66.6% respectively) ($\chi^2=19.8$, $p<0.0001$, $df=2$). Risk analysis conferred a risk of 3 fold for heterozygous genotype ($AG+CG$) when compared to the wild type genotype ($AA+CC$) subjects (OR=3.12, 95%C.I.=1.44-6.7, $p=0.003$, $FDR_p=0.009$). The subjects carrying both the single allelic variant and double allelic variant genotype ($AG+GG$)+($CG+GG$) depicted a significant difference in the frequency distribution between cases (90.7%) and controls (77.0%). Henceforth when compared to the wild type genotype a high risk was observed (OR=2.86, 95%C.I.=1.35-6.07, $p=0.015$, $FDR_p=0.039$) as detailed in table 5.91. Similar to the above combination next genotypic combination of *XRCC1 Pro²⁰⁶Pro* and *MUTYH Gln³²⁴His* modified the lung cancer predisposition. Subjects carrying single allelic genotype ($AG+GC$) for the group showed a high frequency of subjects in cases (82.4%) compared to controls (75.7%) ($\chi^2=12.2$, $p=0.002$, $df=2$). Logistic regression analysis suggested a risk of nearly 3 fold (OR=2.79, 95%C.I.=1.60-4.86, $p=0.0003$, $FDR_p=0.002$) when compared with subjects carrying both wild allele for the two SNPs. For the next genotypic combination of *XRCC1 Pro²⁰⁶Pro* and *XRCC4 intron 3* the frequency distribution revealed there were 66.6% of heterozygous genotype in cases and 54.9 % in controls. Though the difference was not significant a visible trend towards increased risk was observed (OR=1.87, 95%C.I.=0.97-3.42, $p=0.059$). The next genotypic combination of *XRCC1 Pro²⁰⁶Pro* with both *XRCC1 61 C>G* and *XRCC7 6721 G>T* did not show any correlation with lung cancer risk.

The next important genotypic combination that was found to modulate the susceptibility towards lung cancer risk in North Indian population is of *XRCC1 632* and *OGG1 Ser³²⁶Cys*. The frequency of cases with heterozygous genotype ($GA+CG$) was higher as compared to controls (88.7 vs 60.6%) ($\chi^2=23.6$, $p<0.0001$, $df=2$). Therefore, taking wild type genotype ($GG+CC$) as a reference subjects with heterozygous genotype were observed to acquire a high risk towards lung cancer (OR=7.92, 95%C.I.=2.68-23.4, $p=0.0002$, $FDR_p=0.002$). *XRCC1 Gln⁶³²Gln* along with *MUTYH Gln³²⁴His* also effectively modulated lung cancer risk. As shown in table 5.91 an increased frequency of subjects having heterozygous genotype ($GA+GC$) was estimated among cases 67.0% as compared to controls 31.5% ($\chi^2=26.3$, $p<0.0001$, $df=2$). The risk analysis showed a 5-fold risk towards lung cancer (OR=5.07, 95%C.I.=2.66-9.67, $p<0.0001$, $FDR_p=0.002$). For the genotypic combination of *XRCC1 Gln⁶³²Gln* and *XRCC4 Intron3 I/D* there was a higher representation of single allelic carriers ($GA+ID$) in cases compared to controls (59.2 vs 44.3%) ($\chi^2=19.8$, $p<0.0001$, $df=2$). Cox-regression analysis revealed a high death rate (H.R.=4.23, 95% C.I.=1.92-9.30, $p=0.0003$, $FDR_p=0.002$). Subjects carrying mutant genotype ($AA+DD$) for the combination revealed a high death rate (H.R.=2.47, 95% C.I.=1.55-3.9, $p=0.0001$, $FDR_p=0.002$).

Table 5.91 shows a significantly increased frequency of the subjects having heterozygous genotype ($GG+CC$) for the combination of $XRCC1$ $Gln^{632}Gln$ and $XRCC6$ $61\ C>G$ in cases (65.2%) compared to controls (50.0%). Therefore, risk analysis estimated an increased odds ratio favouring the risk of developing lung cancer in patients with heterozygous genotype ($GG+CC$) (OR=1.94, 95%C.I.=1.04-3.62, **$p=0.035$** , $FDR_p=0.07$). The combination of $XRCC1$ $His^{632}His$ with $XRCC7$ $6721\ G>T$ exhibited an overall increased risk towards lung cancer. The frequency distribution suggested that both the heterozygous (61.4% vs 53.54%) and mutant genotype (28.0% vs 20.5%) was higher in cases compared to the controls respectively ($\chi^2=9.27$, **$p=0.0097$** , $df=2$). Furthermore, heterozygous genotype ($GA+GT$) acquired a risk of 2.5 fold (OR=2.54, 95%C.I.=1.16-5.59, **$p=0.019$** , $FDR_p=0.04$). Whereas the risk increased further for the mutant genotype ($AA+TT$) to 3.5 fold (OR=3.5, 95%C.I.=1.42-8.39, **$p=0.006$**) when compared with the wild genotype ($GG+GG$).

The genotypic combination of $XRCC1$ $Arg^{280}His$ and $OGG1$ $Ser^{326}Cys$ revealed a protective effect for the genotypic combinations. Subjects with heterozygous genotype ($GG+CC$) were represented slightly more in cases compared to control subjects (94.0 vs 93.0%). Regression analysis suggested a protective effect towards lung cancer (OR=0.43, 95%C.I.=0.11-1.65, $p=0.22$). No combination of $XRCC1$ $Arg^{280}His$ was associated with lung cancer risk but single allelic carriers for the double combination of $XRCC1$ $Arg^{280}His$ and $XRCC1$ $61C>G$ had a higher representation 94.0% for controls compared to cases (85.1%) and a strong protective effect was observed (OR=0.31, 95%C.I.=0.11-0.85, **$p=0.02$**).

Table 5.91: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk

Genotype	Cases (n%)	Controls (n%)	Overall AOR (95% CI) ^a	p ^b	FDR _p
OGG1 Ser ³²⁶ Cys, C/G XRCC4,I/D	N=140	N=79			
CC/II	8 (5.71)	32 (40.5)	1.00 (Reference)		
CG/ID	122 (87.1)	37 (46.8)	3.7 (1.57-9.0)	0.0030	0.009
GG/DD	10 (7.1)	10 (12.6)	1.02 (0.30-3.39)	0.96	0.96
CG+GG/ID+DD	132 (94.2)	47 (59.4)	3.34 (1.40-7.69)	0.0063	0.02
OGG1 Ser ³²⁶ Cys, C/G XRCC6 61C>G	N=140	N=125			
CC/CC	14 (10.0)	26 (20.8)	1.00 (Reference)		
CG/CG	126 (90.0)	96 (76.8)	1.88 (0.89-3.97)	0.096	0.15
GG/GG	-	3 (2.4)	-	-	-
CG+GG/CG+GG	126 (90.0)	99 (79.2)	1.82 (0.86-3.84)	0.11	0.17
OGG1 Ser ³²⁶ Cys, C/G XRCC7 G>T	N=152	N=132			
CC/GG	9 (5.9)	18 (13.6)	1.00 (Reference)		
CG/GT	135 (88.8)	91 (68.9)	2.47 (1.0-6.05)	0.047	0.09
GG/TT	8 (5.2)	23 (17.4)	1.66 (0.70-3.90)	0.24	0.53
CG+GG/GT/TT	143 (94.0)	114 (86.3)	2.13 (0.86-5.25)	0.09	0.15
Mutyh Gln ³²⁴ His, G/C XRCC4 Intron 3 I/D	N=119	N= 105			
GG/II	42 (35.2)	63 (60.0)	1.00 (Reference)		
GC/ID	76 (63.8)	40 (38.0)	3.05 (1.69-5.5)	0.0002	0.002
CC/DD	1 (0.8)	2 (1.9)	0.86 (0.05-12.53)	0.91	0.96
GC+CC/ID/DD	77 (64.7)	42 (40.0)	2.94 (1.63-5.30)	0.0003	0.002
Mutyh Gln ³²⁴ His, G/C XRCC6 61C>G	N=136	N= 122			
GG/CC	62 (45.5)	81 (66.3)	1.00 (Reference)		
GC/CG	74 (54.4)	39 (31.9)	2.41 (1.39-4.18)	0.0029	0.009
CC/GG	-	2 (1.6)	-	-	
GC+CC/CG+GG	74 (54.4)	41 (33.6)	2.27 (1.32-3.91)	0.003	0.010

to be cont'd on next page

Table 5.91: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk (cont..)

Genotype	Cases (n%)	Controls (n%)	Overall AOR (95% CI) ^a	p ^b	FDR _p
Mutylh Gln ³²⁴ His, G/C XRCC7 G>T	N=112	N= 91			
GG/GG	32 (28.5)	48 (52.7)	1.00 (Reference)		
GC/GT	76 (67.8)	42 (46.1)	2.32 (1.26-4.27)	0.0068	0.019
CC/TT	4 (3.5)	1 (1.0)	4.0 (0.42-39.5)	0.22	0.53
GC+CC/GT/TT	80 (71.4)	43 (47.2)	2.45 (1.36-4.39)	0.0026	0.010
XRCC1 Arg ³⁹⁹ Gln, G/A OGG1 Ser ³²⁶ Cys, C/G	N=86	N=79			
GG/CC	19 (22.0)	32 (42.1)	1.00 (Reference)		
GA/CG	66 (76.7)	37 (46.8)	2.56 (1.21-5.4)	0.013	0.03
AA/GG	1 (1.1)	10 (12.6)	0.32(0.1-0.98)	0.046	0.34
GA+AA/CG/GG	67 (77.9)	47 (59.4)	2.07 (0.99-4.33)	0.05	0.1
XRCC1 Arg ³⁹⁹ Gln, G/A Mutylh Gln ³²⁴ His, G/C	N=123	N= 137			
GG/GG	83 (67.4)	111 (81.0)	1.00 (Reference)		
GA/GC	38 (30.8)	24 (17.5)	2.03(1.11-3.72)	0.021	0.04
AA/CC	2 (1.6)	2 (1.4)	1.18(0.43-3.22)	0.74	0.96
GA+AA/GC+CC	40 (32.5)	26 (18.9)	2.00 (1.1-3.6)	0.02	0.05
XRCC1 Arg ³⁹⁹ Gln, G/A XRCC4 Intron 3 I/D	N=95	N= 88			
GG/II	44 (46.3)	47 (53.4)	1.00 (Reference)		
GA/ID	41 (43.1)	27 (30.6)	1.64 (0.83-3.2)	0.15	0.22
AA/DD	10 (10.5)	14 (15.9)	0.82 (0.51-1.34)	0.44	0.74
GA+AA/ID/DD	51 (53.6)	41 (46.5)	1.31 (0.70-2.4)	0.38	0.47
XRCC1 Arg ³⁹⁹ Gln, G/A XRCC6 61C>G	N=111	N= 94			
GG/CC	71 (63.9)	65 (69.1)	1.00 (Reference)		
GA/CG	39 (35.1)	32 (34.0)	0.94(0.50-1.74)	0.85	0.86
AA/GG	1 (0.9)	2 (2.1)	0.39 (0.10-1.43)	0.15	0.44
GA+AA/CG+GG	40 (36.0)	34 (36.1)	0.88 (0.48-1.61)	0.68	0.13

to be cont'd on next page

Table 5.91: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk (cont..)

Genotype	Cases (n%)	Controls (n%)	Overall AOR (95% CI) ^a	p ^b	FDR _p
XRCC1 Arg ³⁹⁹ Gln, G/A XRCC7 G>T	N=87	N= 71			
GG/GG	31 (35.6)	31 (43.6)	1.00 (Reference)		
GA/GT	44 (50.5)	28 (39.4)	1.33(0.64-2.77)	0.43	0.53
AA/TT	12 (13.7)	12 (16.9)	0.96 (0.59-1.56)	0.89	0.96
GA+AA/GT+TT	56 (64.3)	40 (56.3)	1.23(0.62-2.43)	0.54	0.59
XRCC1 ¹⁹⁴ , C/T OGG1 Ser ³²⁶ Cys, C/G	N= 84	N= 79			
CC/CC	27 (32.1)	49 (62.0)	1 (Reference)		
CT/CG	57 (67.8)	30 (37.9)	3.07 (1.57-5.9)	0.0010	0.004
XRCC1 ¹⁹⁴ , C/T MUTYH Gln ³²⁴ His, G/C	N= 151	N= 188			
CC/GG	116 (8.0)	171 (9.6)	1 (Reference)		
CT/GC	35 (82.4)	17 (75.7)	3.13 (1.61-6.08)	0.0007	0.004
XRCC1 ¹⁹⁴ , C/T XRCC4, I/D	N= 90	N= 102			
CC/II	59 (65.5)	76 (74.5)	1 (Reference)		
CT/ID	31 (34.4)	26 (25.4)	1.45(0.74-2.82)	0.27	0.36
XRCC1 ¹⁹⁴ , C/T XRCC6 61C>G	N= 136	N= 145			
CC/CC	102 (75.0)	116 (80.0)	1 (Reference)		
CT/CG	34 (25.0)	29 (20.0)	1.14(0.62-2.08)	0.65	0.7
XRCC1 ¹⁹⁴ , C/T XRCC7 G>T	N= 86	N= 88			
CC/GG	52 (60.4)	61 (69.3)	1 (Reference)		
CT/GT	34 (39.5)	27 (30.6)	1.25(0.65-2.40)	0.49	0.55
XRCC1 Pro ²⁰⁶ Pro, A/G OGG1 Ser ³²⁶ Cys, C/G	N=140	N=96			
AA/CC	13 (9.2)	22 (22.9)	1 (Reference)		
AG/CG	125 (89.2)	64 (66.6)	3.12 (1.44-6.7)	0.003	0.009
GG/GG	2 (1.42)	10 (10.4)	0.28 (0.05-1.61)	0.15	0.44
AG+GG/CG+GG	127 (90.7)	74 (77.0)	2.86 (1.35-6.07)	0.015	0.039

to be cont'd on next page

Table 5.91: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk (cont..)

Genotype	Cases (n%)	Controls (n%)	Overall AOR (95% CI) ^a	p ^b	FDR _p
XRCC1 Pro ²⁰⁶ Pro, A/G MUTYH Gln ³²⁴ His, G/C	N=119	N= 131			
AA/GG	56 (47.0)	89 (9.6)	1 (Reference)		
AG/GC	63 (82.4)	41 (75.7)	2.79 (1.60-4.86)	0.0003	0.002
GG/CC	-	1 (14.5)	-	-	
AG+GG/GC+CC	63 (92)	42 (41.4)	2.70 (1.56-4.70)	0.0004	0.002
XRCC1 Pro ²⁰⁶ Pro, A/G XRCC4, I/D	N= 112	N= 102			
AA/II	30 (26.7)	38 (37.2)	1 (Reference)		
AG/ID	75 (66.6)	56 (54.9)	1.87(0.97-3.4)	0.05	0.09
GG/DD	7 (6.2)	8 (7.8)	1.16(0.34-3.8)	0.80	0.96
AG+GG/ID+DD	82 (73.2)	64 (62.7)	1.54 (0.83-2.8)	0.16	0.21
XRCC1 Pro ²⁰⁶ Pro, A/G XRCC6 61C>G	N=121	N= 131			
AA/CC	49 (40.4)	62 (47.3)	1 (Reference)		
AG/CG	72 (59.5)	66 (50.3)	1.23 (0.73-2.08)	0.42	0.53
GG/GG	-	3 (2.2)	-	-	
AG+GG/CG+GG	72 (59.5)	69 (52.6)	1.23 (0.73-2.08)	0.42	0.50
XRCC1 Pro ²⁰⁶ Pro, A/G XRCC7 G>T	N=111	N= 111			
AA/GG	26 (23.4)	29 (26.1)	1 (Reference)		
AG/GT	77 (69.3)	67 (60.3)	1.25 (0.65-2.39)	0.49	0.55
GG/TT	8 (7.2)	15 (13.5)	0.70 (0.23-2.07)	0.61	0.96
AG+GG/GT+TT	85 (76.5)	82 (73.8)	1.17 (0.61-2.2)	0.62 ^c	0.64
XRCC1 Gln ⁶³² Gln, G/A OGG1 Ser ³²⁶ Cys, C/G	N= 116	N= 94			
GG/CC	5 (4.3)	21 (22.3)	1 (Reference)		
GA/CG	103 (88.7)	57 (60.6)	7.92 (2.68-23.4)	0.0002	0.002
AA/GG	8 (7.2)	16 (17.0)	1.77 (0.44-7.008)	0.41	0.74
GA+AA/CG+GG	111 (95.6)	73 (77.6)	7.05 (2.39-20.75)	0.0004	0.02

to be cont'd on next page

Table 5.91: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk (cont..)

Genotype	Cases (n%)	Controls (n%)	Overall AOR (95% CI) ^a	p ^b	FDR _p
XRCC1 Gln ⁶³² Gln, G/A MUTYH Gln ³²⁴ His, G/C	N= 91	N= 114			
GG/GG	28 (30.7)	76 (66.6)	1 (Reference)		
GA/GC	61 (67.0)	36 (31.5)	5.07 (2.66-9.67)	< 0.0001	0.002
AA/CC	2 (2.1)	2 (1.7)	2.93 (0.71-12.06)	0.13	0.44
GA+AA/GC+CC	63 (69.2)	38 (33.3)	4.94 (2.64-9.23)	< 0.0001	0.0013
XRCC1 Gln ⁶³² Gln, G/A XRCC4, I/D	N= 103	N= 106			
GG/II	11 (10.6)	39 (36.7)	1 (Reference)		
GA/ID	61 (59.2)	47 (44.3)	4.23 (1.92-9.30)	0.0003	0.002
AA/DD	31 (30.0)	20 (18.8)	2.47 (1.55-3.9)	0.0001	0.002
GA+AA/ID+DD	92 (89.3)	67 (63.2)	4.77(2.2-10.1)	< 0.0001	0.0013
XRCC1 Gln ⁶³² Gln, G/A XRCC6 61C>G	N= 95	N= 114			
GG/CC	26 (27.3)	52 (45.6)	1 (Reference)		
GA/CG	62 (65.2)	57 (50.0)	1.94 (1.04-3.62)	0.035	0.07
AA/GG	7 (7.3)	5 (4.3)	2.65 (0.70-10.03)	0.14	0.44
GA+AA/CG+GG	69 (72.6)	62 (54.3)	2.00 (1.09-3.69)	0.025	0.05
XRCC1 Gln ⁶³² Gln, G/A XRCC7 G>T	N= 114	N= 112			
GG/GG	12 (10.5)	29 (25.8)	1 (Reference)		
GA/GT	70 (61.4)	60 (53.5)	2.54 (1.16-5.59)	0.019	0.04
AA/TT	32 (28.0)	23 (20.5)	3.46 (1.42-8.39)	0.006	0.07
GA+AA/GT+TT	102 (89.4)	83 (74.1)	2.80 (1.31-5.97)	0.007	0.020
XRCC1 Arg ²⁸⁰ His, G/A OGG1 Ser ³²⁶ Cys, C/G	N= 202	N= 158			
GG/CC	9 (4.4)	3 (1.8)	1 (Reference)		
GA/CG	190 (94.0)	147 (93.0)	0.43 (0.11-1.65)	0.22	0.3
AA/GG	3 (1.4)	8 (8.0)	0.20(0.02-1.9)	0.16	0.44
GA+AA/CG+GG	193 (95.5)	155 (98.1)	0.71 (0.24-2.1)	0.54	0.59

to be cont'd on next page

Table 5.91: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk (cont..)

Genotype	Cases (n%)	Controls (n%)	Overall AOR (95% CI) ^a	p ^b	FDR _p
XRCC1 Arg ²⁸⁰ His, G/A MUTYH Gln ³²⁴ His, G/C	N= 119	N= 86			
GG/GG	13 (25.2)	16 (18.6)	1 (Reference)		
GA/GC	105 (88.2)	68 (79.0)	2.08(0.90-4.80)	0.08	0.13
AA+CC	1 (0.8)	2 (2.3)	0.69 (0.05-9.03)	0.78	0.96
GA+AA/GC+CC	106 (89.0)	70 (81.3)	2.03 (0.88-4.67)	0.09	0.15
XRCC1 Arg ²⁸⁰ His, G/A XRCC4, I/D	N= 137	N= 123			
GG/II	10 (7.2)	9 (7.3)	1 (Reference)		
GA/ID	113 (82.4)	108 (87.8)	0.91(0.34-2.43)	0.86	0.86
AA/DD	14 (10.2)	6 (4.8)	1.88(0.44-7.94)	0.38	0.74
GA+AA/ID+DD	127 (92.7)	114 (92.6)	0.97 (0.37-2.58)	0.96	0.96
XRCC1 Arg ²⁸⁰ His, G/A ,XRCC6 61C>G	N= 135	N= 135			
GG/CC	17 (12.5)	7 (5.1)	1 (Reference)		
GA/CG	115 (85.1)	127 (94.0)	0.31(0.11-0.85)	0.02	0.04
AA/GG	3 (2.2)	1 (0.7)	1.06 (0.08-14.2)	0.95	0.96
GA+AA/CG+GG	118 (87.4)	128(94.8)	0.51 (0.22-1.19)	0.12	0.17
XRCC1 Arg ²⁸⁰ His, G/A XRCC7 G>T	N= 128	N= 133			
GG/GG	5 (3.9)	3 (2.2)	1 (Reference)		
GA/GT	114 (89.0)	122 (91.7)	0.60 (0.13-2.65)	0.5	0.56
AA/TT	9 (7.0)	8 (6.0)	0.77 (0.13-4.54)	0.77	0.96
GA+AA/GT+TT	123 (96.0)	130 (97.7)	1.26 (0.78-6.5)	0.13	0.18

(a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. (b)-Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 3: combined hetero and mutant genotype 4:-FDR - False Discovery Rate Adjusted-p

5.14.3 MDR analysis of BER and DSB repair pathway genes

Table 5.92 summarizes the MDR results for the BER (Base excision repair) and DSB (Double strand break repair) pathway genes. Gene-gene and gene-environment interaction has been demonstrated in the table. All the possible interactions model has listed in the table for lung cancer patients. Stratified analysis was done for the various histological subgroups and smoking status. Out of all interaction models, those having the maximum cross validation consistency (CVC), minimum prediction error (1-testing accuracy) and permutation p-value was considered as the best model. Overall the best interaction model was the three factor model *XRCC1 206*, *XRCC1 632*, *OGG1 Ser³²⁶Cys* which has a CVC of 10/10, prediction error=0.33 and ***p*=0.0014** as shown in table 5.92. The second best model was four factor model comprising *XRCC1 206*, *XRCC1 632*, *XRCC1 280*, *OGG1 Ser³²⁶Cys* with CVC of 9/10 and prediction error of 0.29 and *p*=0.0008 as detailed in table 5.92. The interaction model suggests that the four genes *i.e.* *XRCC1 206*, *XRCC1 632*, *XRCC1 280* and *OGG1 Ser³²⁶Cys* significantly contributes in governing lung cancer predisposition. Further, interaction analysis was attempted was between the variants of these genes under different categories so as to find out best interacting group and therefore contributing factors. In smokers, the best interaction model was the three factor model consisting of *XRCC1 206*, *632* and *OGG1 Ser³²⁶Cys* as shown in table 5.92. The CVC for the model was 9/10 with prediction error 0.36 and *p*=0.05. Similarly, the best interaction model in the category of non-smokers was of three factor model *XRCC1 206*, *632* and *OGG1 Ser³²⁶Cys* with a CVC of 10/10 prediction error 0.20 and ***p*-value=0.02**. The best interacting models were also analyzed for the three histology. It was observed in case of subjects diagnosed with ADCC subtype the three factor model was the best model *XRCC1 206*, *632* and *OGG1 Ser³²⁶Cys*. The model showed a CVC OF 10/10, with the least prediction error (0.29) and high *p*-value=0.01. In case of SQCC, four factor model was the best model with CVC of 9/10 and prediction error of 0.32 and significant *p*-value=0.02. Furthermore, similar to ADCC subtype for subjects with SCLC subtype, three factor model was the best interacting model with three factors *XRCC1 206*, *632* and *OGG1 Ser³²⁶Cys* contributing to the lung cancer predisposition as table 5.92 indicated.

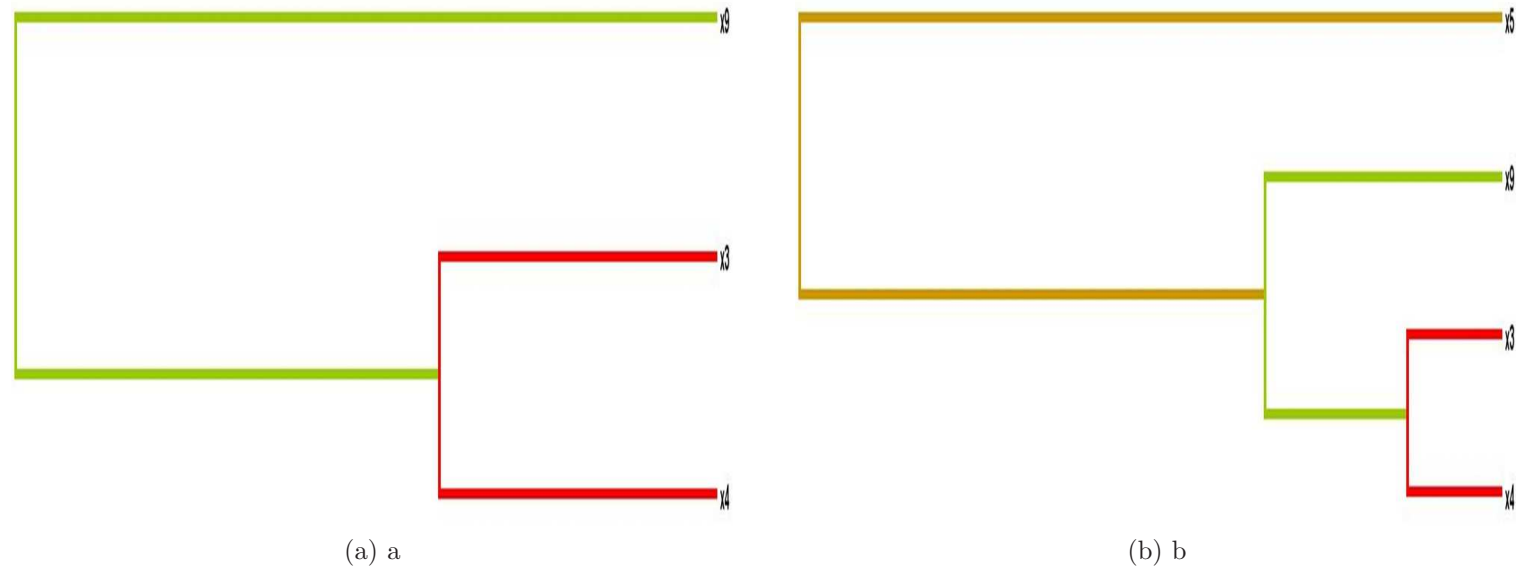


Figure 5.30: Multifactor dimensionality reduction (MDR) interaction dendrogram between BER (*OGG1*, *MUTYH*, *XRCC1*) and DSB repair genes (*XRCC4*, *XRCC6*, *XRCC7*) X1-*XRCC1* 399, X2-*XRCC1* 194, X3-*XRCC1* 206, X4-*XRCC1* 632, X5-*XRCC1* 280, X6-*XRCC6*, X7-*XRCC7*, X8-*MUTYH*, X9-*OGG1*, X10-*XRCC4* (a) In all cases and controls best three factor model (b) In all cases and controls best four factor model

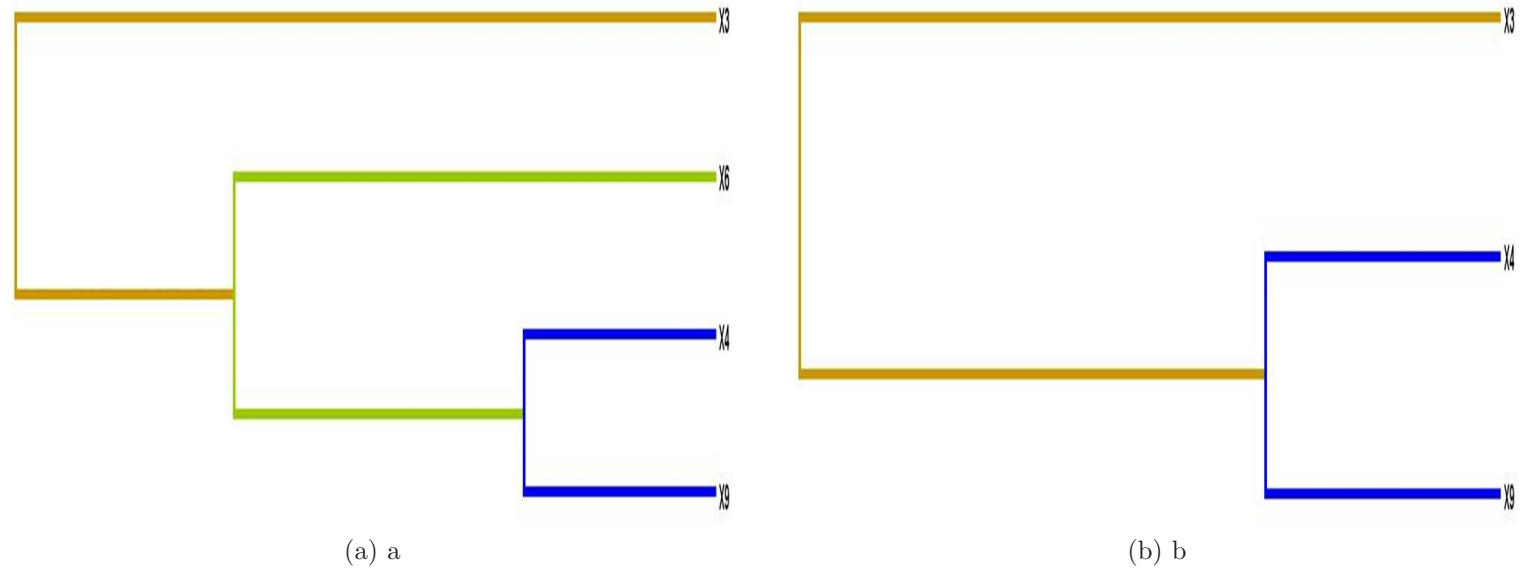


Figure 5.31: Multifactor dimensionality reduction (MDR) interaction dendrogram between BER (*OGG1*, *MUTYH*, *XRCC1*) and DSB repair genes (*XRCC4*, *XRCC6*, *XRCC7*): X1-*XRCC1* 399, X2-*XRCC1* 194, X3-*XRCC1* 206, X4-*XRCC1* 632, X5-*XRCC1* 280, X6-*XRCC6*, X7-*XRCC7*, X8-*MUTYH*, X9-*OGG1*, X10-*XRCC4* (a) For four factor model in ADCC subjects (b) For three factor model in ADCC subjects

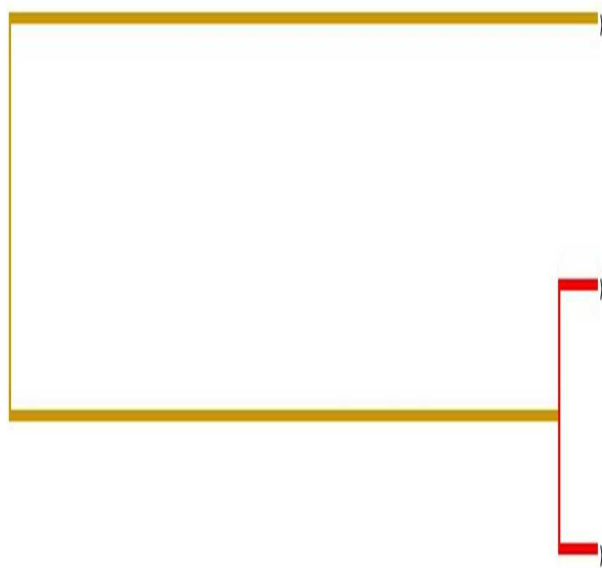


Figure 5.32: Multifactor dimensionality reduction (MDR) interaction dendrogram between BER (*OGG1*, *MUTYH*, *XRCC1*) and DSB repair genes (*XRCC4*, *XRCC6*, *XRCC7*): X1-XRCC1 399, X2-XRCC1 194, X3-XRCC1 206, X4-XRCC1 632, X5-XRCC1 280, X6-XRCC6, X7-XRCC7, X8-MUTYH, X9-OGG1, X10-XRCC4 in smokers

The entropy dendrogram for the two excision repair pathway genes is mentioned in figure 5.30. Entropy dendrogram for smokers is detailed in figure 5.32 and for histology the figure 5.31 showcased the entropy dendrogram. The dendrogram indicate the extent of interaction of interacting SNPs. The length of the lines suggests the intensity of the interaction. The longer the lines weaker is the interaction. In addition to this, the degree of interaction is depicted by the color of the lines in the dendrogram. Red and Orange indicate synergy. Yellow suggests independence, whereas green and blue indicate no-interaction.

Table 5.92: Multifactor Dimensionality reduction (MDR) analysis showing association of high order interactions with lung cancer

No. of risk factors	Best interaction Model	Cross Validation Consistency (CVC)	Prediction Error	p-value
1	OGG1 326	9/10	0.41	0.09
2	XRCC1 632, OGG1 326	9/10	0.36	0.02
3	XRCC1 206, XRCC1 632, OGG1 326	10/10	0.33	0.0014
4	XRCC1 206, XRCC1 632, XRCC1 280, OGG1 326	9/10	0.29	0.0008

to be cont'd on next page

Table 5.92: Multifactor Dimensionality reduction (MDR) analysis showing association of high order interactions with lung cancer (cont..)

No. of risk factors	Best interaction Model	Cross Validation Consistency (CVC)	Prediction Error	p-value
5	XRCC1 399, XRCC1 206, XRCC1 632, OGG1 326, XRCC4	6/10	0.37	0.02
MDR analysis showing interactions of BER and DSB genes with lung cancer risk in smokers				
1	XRCC1 632	8/10	0.41	0.14
2	XRCC1 206, XRCC1 632	7/10	0.37	0.07
3	XRCC1 206, XRCC1 632, OGG1 326	9/10	0.36	0.05
MDR analysis showing interactions of BER and DSB genes with lung cancer risk in non-smokers				
3	XRCC1 206, XRCC1 632, OGG1 326	10/10	0.20	0.02
MDR analysis showing interactions of BER and DSB genes with lung cancer risk in ADCC subjects				
1	OGG1 Ser ³²⁶ Cys	10/10	0.35	0.07
2	XRCC1 632, OGG1 326	10/10	0.34	0.06
3	XRCC1 206, XRCC1 632, OGG1 326	10/10	0.29	0.01
4	XRCC1 206, XRCC1 632, XRCC6 C61G, OGG1 326	9/10	0.30	0.02
MDR analysis showing interactions of BER and DSB genes with lung cancer risk in SQCC subjects				
3	XRCC1 206, XRCC1 632, MUTYH 324	8/10	0.33	0.03
4	XRCC1 206, XRCC1 632, MUTYH 324, OGG1 326	9/10	0.32	0.02
6	XRCC1 399, XRCC1 206, XRCC1 632, MUTYH 324, OGG1 326, XRCC4	9/10	0.32	0.02
MDR analysis showing interactions of BER and DSB genes with lung cancer risk in SCLC subjects				
2	OGG1 326, XRCC1 632	6/10	0.42	0.39
4	XRCC1 206, XRCC1 632, OGG1 326	10/10	0.36	0.17

5.14.4 CART analysis of BER and DSB genes

CART analysis, is a tree building technique, that utilizes binary recursive partitioning to find the different genotypic groups affecting lung cancer predisposition, which is not possible to obtain from traditional logistic regression analysis. A large number of variables are taken into account and high risk possessing subgroups are identified. The outcome of

the analysis is a decision tree which is a result of data mining approach utilized in the analysis. The decision tree depicts various factors and their interaction and risk associated with the combination. The most significant factor acts as the root split. Whereas, the terminal nodes are the high risk posing subgroups. Terminal node with least case rate is considered the reference. Odds ratio is respectively calculated.

Table 5.93: Overall Risk estimates based on CART analysis (BER and DSBH genes)

Node	Genotype in each terminal node	Control (n)	Case (n)	Case Rate ^a	OR (95% CI)	p-value ^b
Node 1	XRCC1 632(H+M)/ MUTYH 324(H+M)/ XRCC1 206(W)	16	62	79.0	20.9 (45.5-9.6)	0.0001
Node 2	XRCC1 632(H+M)/ MUTYH 324(H+M)/ XRCC1 206(H+M)/ XRCC1 280(W)	0	9	100	0.587789e+8 (-0.000)	0.97
Node 3	XRCC1 632(H+M)/ MUTYH 324(H+M)/ XRCC1 206(H+M)/ XRCC1 280(H+M)/ XRCC6 61C>G(H+M) / XRCC1 194(W)	14	33	70.2	12.7 (29.2-5.53)	0.0001
Node 4	XRCC1 632(H+M)/ MUTYH 324(H+M)/ XRCC1 206(H+M)/ XRCC1 280(H+M)/ XRCC6 61C>G(H+M)/ XRCC1 194(H+M)	5	3	37.5	3.24 (15.0-0.69)	0.133
Node 5	XRCC1 632(H+M)/ MUTYH 324(H+M)/ XRCC1 206(H+M)/ XRCC1 280(H+M)/ XRCC6 61C>G(W)	25	20	44.4	4.32 (9.6-1.92)	0.00037
Node 6	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(H+M)	14	28	66.6	10.8(25.1-4.6)	0.0001
Node 7	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(H+M)/ OGG1 326(H+M)/ XRCC1 399(H+M) / XRCC4(W)	16	22	57.8	7.42 (17.3-3.18)	0.0001
Node 8	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(H+M)/ OGG1 326(H+M)/ XRCC1 399(H+M) / XRCC4(H+M) /XRCC1 280(W)	0	1	100	0.587789e+8 (0.0-0.00)	0.99
Node 9	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(H+M)/ OGG1 326(H+M)/ XRCC1 399(H+M) / XRCC4(H+M) /XRCC1 280(H+M)/ XRCC6 61C>G(W)	9	15	62.5	9.00 (24.2-3.33)	0.00001
Node 10	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(H+M)/ OGG1 326(H+M)/ XRCC1 399(H+M) / XRCC4(H+M) /XRCC1 280(H+M)/ XRCC6 61C>G(W)/ XRCC1 206(W)	8	5	38.4	3.37 (11.7-0.97)	0.055
Node 11	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(H+M)/ OGG1 326(H+M)/ XRCC1 399(H+M) / XRCC4(H+M) /XRCC1 280(H+M)/ XRCC6 61C>G(H+M))	29	23	44.2	4.28 (9.3-1.97)	0.00024
Node 12	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(W)/ OGG1 326(H+M)/ XRCC1 399(W) / XRCC7 6721(H+M)/ XRCC4(W)	2	6	75.0	16.2 (88.0-2.9)	0.00126

to be cont'd on next page

Table 5.93: Overall Risk estimates based on CART analysis (BER and DSBH genes) (cont..)

Node	Genotype in each terminal node	Control (n)	Case (n)	Case Rate ^a	OR (95% CI)	p-value ^b
Node 13	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(W)/ OGG1 326(H+M)/ XRCC1 399(W) / XRCC7 6721(H+M)/ XRCC4(H+M)	14	10	41.6	3.85 (10.2-1.44)	0.0069
Node 14	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(W)/ OGG1 326(H+M)/ XRCC1 399(W) / XRCC7 6721(W)	10	2	16.7	1.0 (5.4-0.21)	0.92
Node 15	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(W)/ OGG1 326(W)/ XRCC1 280(W)	0	5	100	0.587789e+8 (0.0-0.00)	0.984
Node 16	XRCC1 632(H+M)/ MUTYH 324(W)/ XRCC1 194(W)/ OGG1 326(W)/ XRCC1 280(H+M)	23	4	17.8	0.93 (3.1-0.28)	0.91
Node 17	XRCC1 632(W)/ XRCC1 206(H+M)/ XRCC1 280(H+M)/ OGG1 326(H+M)/ XRCC1 194(W)	17	35	67.3	11.1 (24.7-4.99)	0.001
Node 18	XRCC1 632(W)/ XRCC1 206(H+M)/ XRCC1 280(H+M)/ OGG1 326(H+M)/ XRCC1 194(H+M)	6	3	33.3	2.7 (11.9-0.60)	0.19
Node 19	XRCC1 632(W)/ XRCC1 206(H+M)/ XRCC1 280(H+M)/ OGG1 326(W)/	5	1	16.6	1.0 (9.9-0.11)	0.94
Node 20	XRCC1 632(W)/ XRCC1 206(H+M)/ XRCC1 280(W)/	8	0	0		-
Node 21	XRCC1 632(W)/ XRCC1 206(W)	81	15	15.6	Reference	-

(a)-case rate is percentage LC cases among all individuals in each terminal node case/(case+ control) x100;(b)-p<0.05 is considered statistically significant; (c) reference group; adjustment for age, gender and pack years is done

CART analysis was performed for the DNA repair pathway genes as shown in figure. The tree shows the initial split was formed by *XRCC1 632*. Therefore the SNP is the most important factor among the interacting factors as shown in table 5.93. Table 5.93 shows the odds ratio, p-values of interacting subgroups. Terminal node 21 with the genotypic combination of *XRCC1 632* and *XRCC1 206* had the lowest case rate of 15.6%. It is therefore considered as reference (figure 5.33). Individuals carrying the mutant genotype for *XRCC1 632(M)*, *MUTYH 324(M)* and *XRCC1 206(W)* i.e. Node 1 conferred the highest risk towards lung cancer (OR=20.9, 95%C.I.=45.5-9.6, ***p*=0.00001**). Terminal node 12 with the genetic combination of *XRCC1 632(M)*, *MUTYH 324(W)*, *XRCC1 194(W)*, *OGG1 326(M)*, *XRCC1 399(W)*, *XRCC7 6721(M)*, *XRCC4(W)* also possessed a significantly high risk towards lung cancer (OR=16.2, 95%C.I.=88.0-2.9, ***p*=0.0013**). Subjects with genotypic combination of node 3 carrying *XRCC1 632(H+M)*/*MUTYH 324(H+M)*/*XRCC1 206 (H+M)*/*XRCC1 280 (H+M)*/*XRCC6 61 C>GH+M*/*XRCC1 194(W)* exhibited an odds of twelve folds (OR=12.7, 95% C.I.=29.2-5.53, ***p*=0.0001**). Node6 showed an almost eleven fold risk carrying the genotypic combination (*XRCC1 632(H+M)*/*MUTYH 324(W)*/*XRCC1 194(H+M)*) (OR=10.8, 95%C.I.=25.1-4.6, ***p*=0.0001**). Similarly, node 7 carrying the genotypic combination (*XRCC1 632(H+M)*/*MUTYH 324(W)*/*XRCC1 194(+M)*/*OGG1 326(H+M)*/*XRCC1 399(H+M)*/*XRCC4(W)*) showed a high lung cancer risk (OR=7.42, 95% C.I.=17.3-3.18, ***p*=0.0001**). Further, subjects under with the genotypic combination of terminal node 16 and 19 did not show any association with lung cancer as shown in table 5.23.

Further, to explore gene-environment interactions CART analysis incorporating smoking status was also performed. Figure 5.34 depicted the respective tree structure. The initial split was in smoking indicating smoking act as an important factor in lung carcinogenesis. The terminal nodes were different for smokers and non-smokers. The table 5.95 shows the data for 16 nodes. Further inspection of the CART structure suggested distinct risk genotypic combinations for smokers and non-smokers. The least case rate was obtained for terminal node 16 with a case rate of 14.7% and it was observed as a reference. The genotype for terminal node 16 was *OGG1 326(W)*. In the group of smokers, the most significant factor was *XRCC1 632*. Subgroups under smokers category with the highest risk was of terminal node 5. The genotypic combination for the node was *XRCC1 632(M)*, *MUTYH 324(M)*, *XRCC1 194(W)*, *XRCC4(M)*, *XRCC1 280(W)*. Terminal node5 with genotypic combination *XRCC1 632(M)*, *MUTYH 324(M)*, *XRCC1 194(W)*, *XRCC4(M)*, *XRCC1 280(W)* showed the highest risk towards lung cancer with an AOR of 22.4 (OR=22.4, 95%C.I.=31.9-15.7, ***p*=0.0004**). Node 2 with the genotypic combination of *XRCC1 632(M)*, *MUTYH 324(W)*, *XRCC1 194(M)* possessed a significant high risk (OR=12.7, 95%C.I.=14.9-10.7, ***p*=0.0001**). Similarly nodes 1, 3, 7, 9

were high risk subgroups as shown in table 5.95. In non-smokers, *OGG1 326* acted as a significant risk factor. Terminal node 13 possessed a very high risk of 30 fold towards lung cancer predisposition. The genotype of node 13 was *OGG1 326(M)*, *MUTYH 324(M)*, *XRCC1 C61G(M)*, *XRCC1 399(W)* (OR=30.0, 95%C.I.=44.5-20.9, ***p=0.0001***). Remaining three subgroups under non-smokers showed a high risk towards lung cancer as shown in table 5.95. The genotypic groups under non-smokers had a combination of *OGG1*, *MUTYH*, *XRCC6* and *XRCC1 399* indicating their involvement in lung cancer for non-smokers.

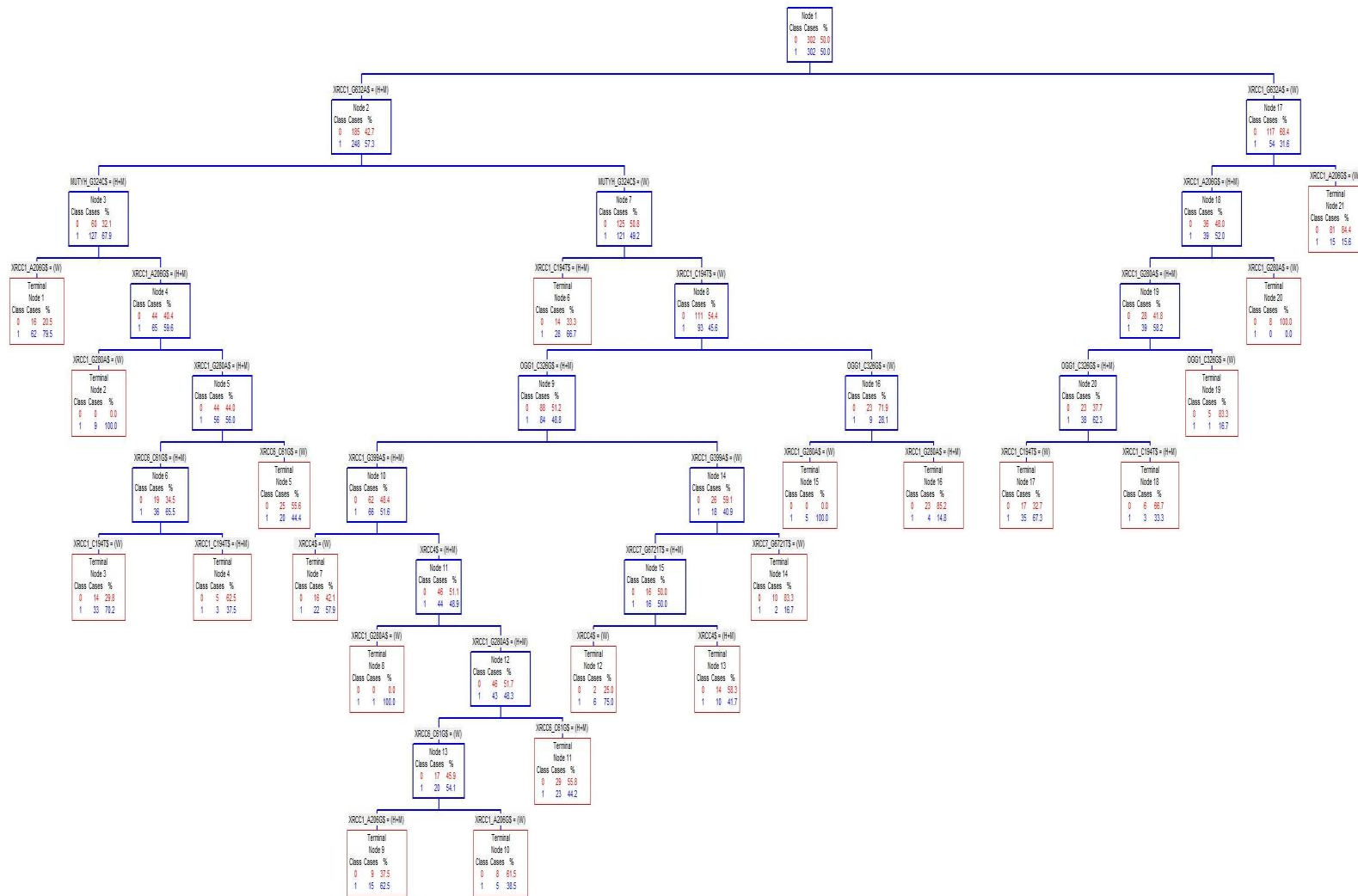


Figure 5.33: CART analysis in polymorphic DNA repair pathway genes. W= homozygous wild type genotype, H+M= heterozygous + homozygous variant genotype

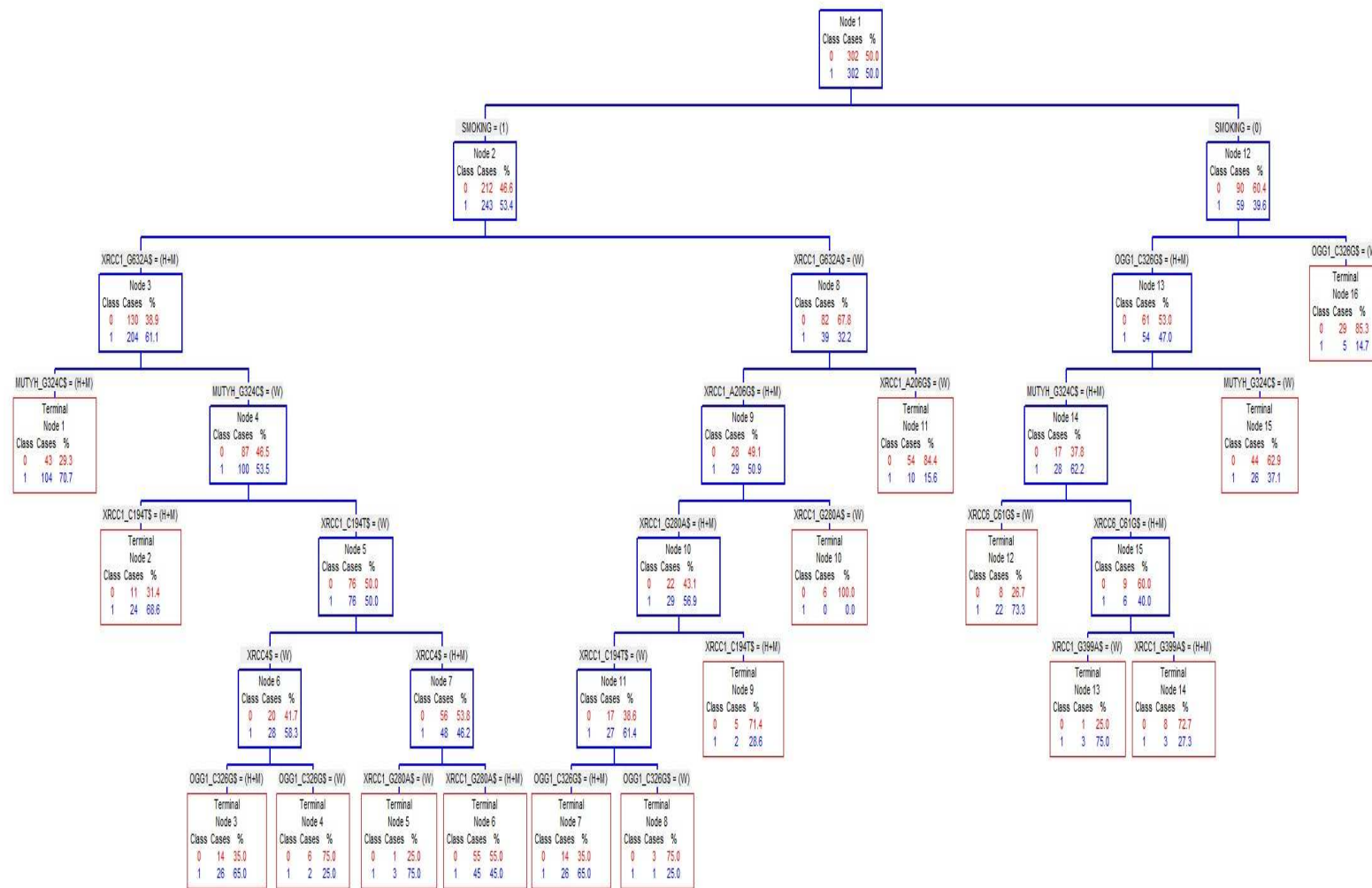


Figure 5.34: CART analysis in polymorphic DNA repair pathway genes. W= homozygous wild type genotype, H+M= heterozygous + homozygous variant genotype

Table 5.94: Overall risk estimates based on CART analysis (BER and NER genes) among smokers

Terminal Node in smokers	Genotype in each terminal node	Control (n)	Case (n)	Case Rate ^a	OR (95% CI)	p-value ^b
Node 1	XRCC1 632(M)/ MUTYH 324(M)	43	104	70.7	15.4 (17.8-13.4)	0.0001
Node 2	XRCC1 632(M)/ MUTYH 324(W)/ XRCC1 194(M)	11	24	68.5	12.6 (14.9-10.7)	0.0001
Node 3	XRCC1 632(M)/ MUTYH 324(M)/ XRCC1 194(W)/ XRCC4(W)/ OGG1 326(M)	14	26	65.0	11.9 (14.0-10.2)	0.0001
Node 4	XRCC1 632(M)/ MUTYH 324(M)/ XRCC1 194(W)/ XRCC4(W)/ OGG1 326(W)	6	2	25.0	1.82 (2.3-1.4)	0.133
Node 5	XRCC1 632(M)/ MUTYH 324(M)/ XRCC1 194(W)// XRCC4 (M)/ XRCC1 280(W)	1	3	75.0	22.4 (31.9-15.7)	0.0004
Node 6	XRCC1 632(M)/ MUTYH 324(M)/ XRCC1 194(W)// XRCC4 (M)/ XRCC1 280(M)	55	45	45.0	5.0 (5.8-4.4)	0.0001
Node 7	XRCC1 632(W)/ XRCC1 206(M)/ XRCC1 280(M)/ XRCC1 194(W)/ OGG1 326(M)	14	26	65.0	12.2 (14.3-10.4)	0.0001
Node 8	XRCC1 632(W)/ XRCC1 206(M)/ XRCC1 280(M)/ XRCC1 194(W)/ OGG1 326(W)	0	1	100.0	1.3 (1.9-0.9)	0.14
Node 9	XRCC1 632(W)/ XRCC1 206(M)/ XRCC1 280(M)/ XRCC1 194(M)	5	2	28.5	2.7 (3.5-2.1)	0.00001
Node 10	XRCC1 632(M)/ XRCC1 206(M)/ XRCC1 280(W)	6	0	0	0.00 (-0.00)	0.88
Node 11	XRCC1 632(W)/ XRCC1 206(W)	54	10	15.6	1.07 (1.3-0.9)	0.42

(a)-case rate is percentage LC cases among all individuals in each terminal node case/(case+ control) x100; (b)-p<0.05 is considered statistically significant; c reference group; no adjustment for age, gender and pack years is done

Table 5.95: Overall risk estimates based on CART analysis (BER and NER genes) among non-smokers

Terminal Node in non-smokers	Genotype in each terminal node	Control (n)	Case (n)	Case Rate ^a	OR (95% CI)	p-value ^b
Node 12	OGG1 326(M)/ MUTYH 324(M)/ XRCC6 61(W)	8	22	73.3	18.3 (21.9-15.2)	0.0001
Node 13	OGG1 326(M)/ MUTYH 324(M)/ XRCC6 61 (M)/ XRCC1 399(W)	1	3	75.0	30.5 (44.5-20.9)	0.0001
Node 14	OGG1 326(M)/ MUTYH 324(M)/ XRCC6 61 (M)/ XRCC1 399(M)	8	3	27.2	2.5 (3.1.2.0)	0.0001
Node 15	OGG1 326(M)/ MUTYH 324(W)	44	26	37.1	3.7 (4.3-3.2)	0.0001
Node 16	OGG1 326(W)	29	5	14.7	1 (Reference)	-

(a)-case rate is percentage LC cases among all individuals in each terminal node case/(case+ control) x100; (b)-p<0.05 is considered statistically significant; c reference group; no adjustment for age, gender and pack years is done

5.14.5 Association of unfavourable genotypic combinations of BER (*OGG1*, *MUTYH*, *XRCC1*) and DSBR (*XRCC3*, *XRCC6*, *XRCC7*) genes with overall survival of lung cancer patients

In order to find the cumulative effect of the polymorphic variants of each gene, different genotypic combination of unfavourable genotypes of BER and DSBR genes were analyzed for predicting the overall survival of lung cancer patients as demonstrated in table re. The study revealed that polymorphic *XRCC1 Arg³⁹⁹Gln* did not show any modulation of lung cancer prognosis for combination with any other genotypic variant in the study. The frequency of single allelic carriers for the combination of *XRCC1 Arg³⁹⁹Gln* with *XRCC6 61 C>G*, *XRCC7 6721 G>T*, *MUTYH Gln³²⁴His*, G/C was higher compared to wild type subjects respectively as indicated in table . In case of polymorphic combination of *XRCC1 Arg³⁹⁹Gln* and *XRCC7 6721 G>T* the frequency of mutant subjects (*AA+TT*) was 9.8% compared to 14.8% of wild type subjects (*GG+GG*). The MST was observed to be better for wild subjects (*GG+GG*) in comparison to the mutant subjects (*AA+TT*) (5.7 vs 6.2, log rank p-value=0.96). Cox regression analysis suggested a trend towards increased death for mutant subjects as 1.5 fold hazard rate was observed (H.R=1.45, 95%C.I.=0.45-3.87, p=0.45) when compared to wild type subjects. For the next combination of *XRCC1 Arg³⁹⁹Gln* and *MUTYH Gln³²⁴His* MST of single allelic carriers was observed to be 5.9 months compared to 12.0 months for wild subjects. The hazard rate was sequentially observed to be higher for single allelic carriers indicating a trend towards increased death rate (H.R.=1.61, 95%C.I.=0.78-3.31, p=0.9).

Furthermore, polymorphic *XRCC1 Arg¹⁹⁴Trp* was analyzed for the combination of unfavourable genotype. Survival analysis did not reveal any significant difference in the median survival time and hazard rate for the combination of *XRCC1 Arg¹⁹⁴Trp* with either of the three polymorphic variant *XRCC6 61 C>G*, *XRCC7 6721 G>T*, *MUTYH Gln³²⁴His* as shown in table 5.96. For the combination of *XRCC1 Arg¹⁹⁴Trp* and *OGG1 Ser³²⁶Cys* a better survival was observed for wild type subjects as subjects with heterozygous genotype were observed to have a smaller MST compared to wild type subjects (5.0 vs 12.0 Log rank p-value=0.25). Further, regression analysis showcased a high hazard rate for heterozygous subjects (H.R=1.68, 95%C.I.=0.86-3.39, p=0.14) when compared to subjects with wild type genotype.

XRCC1 Pro²⁰⁶Pro similar to genotypic combination of polymorphic *XRCC1 Arg¹⁹⁴Trp* did not modulate lung cancer prognosis for the adverse genotypic combination with *XRCC6 61 C>G* and *XRCC7 6721 G>T*. However, for the combination of *XRCC1*

*Pro*²⁰⁶*Pro* with *MUTYH Gln*³²⁴*His* subjects carrying single allelic carriers (*CT+GC*) had a smaller MST compared to wild type subjects (6.1 vs 10.3). The adjusted hazards ratio exhibited a high death rate with nearly approaching significance (H.R=1.73, 95%C.I.=0.97-3.08, p=0.06).

A trend of increased hazard rate was observed for subjects harboring double unfavourable genotype in case of combination of *XRCC1 Gln*⁶³²*Gln* with both *XRCC6 61 C>G* and *XRCC7 6721 G>T*. For the genotypic combination of *XRCC1 Gln*⁶³²*Gln* with *OGG1 Ser*³²⁶*Cys* subjects comprising unfavourable genotype for single allele exhibited a better survival of 6.0 months compared to 2.9 months (log rank p-value=0.12) for subjects lacking any unfavourable genotype. Hazards ratio obtained from univariate analysis (H.R=0.35) decreased further when adjusted for various covariates the death rate significantly decreased (H.R=0.11, 95%C.I.=0.016-0.82, **p=0.03**). Genotypic combination of polymorphic variant *XRCC1 Gln*²⁸⁰*His* with other polymorphic variants revealed a trend of high death rate for heterozygous patients as shown in table 5.96. However, for the genotypic combination of *XRCC1 Gln*²⁸⁰*His* with *OGG1 Ser*³²⁶*Cys* it was observed wild type (*GG+CC*) subjects had a better MST compared to subjects with heterozygous genotype (*GC+CG*) (16.1 vs 6.8, **log rank p-value=0.02**). A high hazard rate of two-fold was observed for single allelic carriers when compared to wild subjects (H.R=2.2, 95%C.I.=0.98-5.0, p=0.05). Several different combinations of unfavourable genotypes were also examined for other SNPs but none of them showed any significant association with prognosis.

Table 5.96: Genotypic combinations and overall survival

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R. ^a	H.R. ^b	p-value	FDR _p
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC6 61 C>G	N=88						
GG/CC	29 (32.9)	9.7		1	1		
GA/CG	59 (67.0)	4.7	0.75	1.07	1.08 (0.63-1.86)	0.76	0.89
XRCC1Arg ³⁹⁹ Gln, G/A & XRCC7 6721 G>T	N=81						
GG/GG	12 (14.8)	6.2		1	1		
GA/GT	61 (75.3)	7.1	0.90	0.32	0.92 (0.45-1.87)	0.83	0.92
AA/TT	8 (9.8)	5.7	0.96	0.97	1.45 (0.54-3.87)	0.45	0.72
GA+AA/GT+TT	69 (85.1)	6.8	0.3	0.75	0.84 (0.43-1.65)	0.63	0.80
XRCC1 Arg ³⁹⁹ Gln, G/A & MUTYH Gln ³²⁴ His, G/C	N=92						
GG/GG	33 (19.1)	6.3		1	1		
GA/GC	57 (57.4)	5.9	0.17	1.37	1.21 (0.71-2.06)	0.46	0.89
AA/CC	2 (23.4)	10.1	0.72	1.28	0.78 (0.26-2.34)	0.66	0.80
GA+AA/GC+CC	59 (80.8)	5.9	0.17	1.37	1.19 (0.70-2.01)	0.51	0.71
XRCC1 Arg ³⁹⁹ Gln, G/A & OGG1 Ser ³²⁶ Cys, C/G	N=101						
GG/CC	11 (10.8)	12.0		1	1		
GA/CG	89 (88.1)	5.9	0.16	1.6	1.61 (0.78-3.31)	0.19	0.57
AA/GG	1 (0.9)	1	0.12	-	-		
GA+AA/CG+GG	90 (89.1)	5.4	0.14	-	-		
XRCC1 Arg ³⁹⁹ Gln, G/A & XRCC3 Thr ²⁴¹ Met, C/T	N=83						
GG/CC	27 (32.5)	7.3		1	1		
GA/CT	55 (66.2)	7.2	0.92	0.97	0.96 (0.55-1.67)	0.89	0.92
AA/TT	1 (1.2)	0	-	-	-	-	
GA+AA/CT+TT	56 (67.4)	6	0.98	0.99	0.99 (0.57-1.72)	0.98	0.98
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC6 61 C>G	N=85						
CC/CC	65 (76.4)	9.4		1	1		

to be cont'd on next page

Table 5.96: Genotypic combinations and overall survival(contd..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R ^a	H.R ^b	p-value	FDR _p
CT/CG	19 (22.3)	5.0	0.4	0.81	0.81 (0.43-1.53)	0.53	0.89
TT/GG	1 (1.1)	-	0.19	-	-	-	
CT+TT/CG+GG	20 (23.5)	12.1	0.30	0.75	0.70 (0.38-1.8)	0.25	0.7
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC7 6721 G>T	N=92						
CC/GG	35 (38.0)	7.2		1	1		
CG/GT	21 (22.8)	5.0	0.8	0.95	1.1 (0.54-2.2)	0.76	0.89
XRCC1 Arg ¹⁹⁴ Trp, C/T & MUTYH Gln ³²⁴ His, G/C	N =96						
CC/GG	79 (7.6)	9.7		1	1		
CT/GC	17 (89.0)	6.1	0.86	1.05	1.17 (0.6-2.29)	0.63	0.89
XRCC1 Arg ¹⁹⁴ Trp, C/T & OGG1 Ser ³²⁶ Cys, C/G	N=56						
CC/CC	23 (41.0)	12		1	1		
CT/CG	33 (58.9)	5.0	0.25	1.38	1.68 (0.86-3.39)	0.14	0.57
XRCC1 Arg ¹⁹⁴ Trp, C/T & XRCC3 Thr ²⁴¹ Met, C/T	N=83						
CC/CC	27 (32.5)	7.3		1	1		
CT/CT	55 (66.2)	7.2	0.92	1.06	1.67 (0.84-3.31)	0.14	0.57
TT/TT	1 (1.2)	-		-	-	-	
XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC6 61 C>G	N=78						
AA/CC	30 (38.4)	10.4		1	1		
AG/CG	48 (61.5)	4.3	0.80	1.06	0.97 (0.56-1.66)	0.91	0.92
XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC7 6721 G>T	N=77						
AA/GG	17 (22.0)	10.3		1	1		
AG/GT	55 (71.4)	6.8	0.87	0.95	0.96 (0.48-1.92)	0.92	0.92
GG/TT	5 (6.4)	9.0	0.82	0.82	0.73 (0.36-1.50)	0.40	0.96
AG+GG/GT+TT	60 (77.9)	7.5	0.82	0.093	0.94 (0.47-1.86)	0.87	0.94
XRCC1 Pro ²⁰⁶ Pro, A/G & MUTYH Gln ³²⁴ His, G/C	N=74						

to be cont'd on next page

Table 5.96: Genotypic combinations and overall survival(contd..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R ^a	H.R ^b	p-value	FDR _p
AA/GG	32 (43.2)	10.3		1	1		
AG/GC	42 (56.7)	6.1	0.07	1.51	1.73(0.97-3.08)	0.06	0.49
XRCC1 Pro ²⁰⁶ Pro, A/G & OGG1 Ser ³²⁶ Cys, C/G	N=90						
AA/CC	9 (10.0)	15.7		1	1		
AG/CG	79 (87.7)	6.4	0.12	1.87	1.72 (0.69-4.27)	0.24	0.66
GG/GG	2 (2.2)	9.7	0.47	1.74	-	-	
AG+GG/CG+GG	81 (90.0)	6.8	0.12	1.88	1.72 (0.69-4.3)	0.23	0.7
XRCC1 Pro ²⁰⁶ Pro, A/G & XRCC3 Thr ²⁴¹ Met, C/T	N=98						
AA/CC	44 (44.8)	8.3		1	1		0.89
AG/CT	54 (54.5)	7.2	0.52	1.14	1.09 (0.67-1.77)	0.72	
XRCC1 Gln ⁶³² Gln, G/A & XRCC6 61 C>G	N=69						
GG/CC	19 (27.5)	5.1		1	1		
GA/CG	43 (62.3)	3.5	0.42	0.80	0.95 (0.46-1.95)	0.9	0.92
AA/GG	7 (10.1)	4.5	0.68	0.81	1.46 (0.49-2.08)	0.96	0.48
GA+AA/CG+GG	50 (72.4)	4.1	0.42	0.80	1.01 (0.62-3.42)	0.38	0.7
XRCC1 Gln ⁶³² Gln, G/A & XRCC7 6721 G>T	N=82						
GG/GG	10 (12.1)	5.8		1	1		
GA/GT	51 (62.1)	7.5	0.94	1.02	1.1 (0.51-2.45)	0.7	0.89
AA/TT	21 (25.6)	5.0	0.34	1.42	1.68 (0.91-3.1)	0.09	0.80
GA+AA/GT/TT	72 (87.8)	6.4	0.79	1.09	1.0 (0.54-2.41)	0.72	0.84
XRCC1 Gln ⁶³² Gln, G/A & MUTYH Gln ³²⁴ His, G/C	N=63						
GG/GG	19 (30.1)	5.7		1	1		
GA/GC	42 (66.6)	5.8	0.31	1.37	1.39 (0.70-2.7)	0.3	0.74
AA/CC	2 (3.1)	10.1	0.64	1.40	1.39 (0.17-11.3)	0.7	0.56
GA+AA/GC+CC	44 (69.8)	5.9	0.30	1.37	1.39 (0.69-2.8)	0.34	0.7

to be cont'd on next page

Table 5.96: Genotypic combinations and overall survival(contd..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R ^a	H.R ^b	p-value	FDR _p
XRCC1 Gln ⁶³² Gln, G/A & OGG1 Ser ³²⁶ Cys, C/G	N=76						
GG/CC	2 (2.6)	2.9		1	1		
GA/CG	69 (90.7)	6.0	0.12	0.35	0.11 (0.016-0.82)	0.03	0.49
AA/GG	5 (6.5)	9.4	0.05	0.25	-	-	
GA+AA/CG+GG	74 (97.3)	6.3	0.10	0.33	0.20 (0.035-1.23)	0.08	0.7
XRCC1 Gln ⁶³² Gln, G/A & XRCC3 Thr ²⁴¹ Met, C/T	N=74						
GG/CC	23 (31.0)	7.1		1	1		
GA/CT	48 (64.8)	6.3	0.44	1.22	1.16 (0.63-2.13)	0.61	0.89
AA/TT	3 (4.0)	-	-	0-	-	-	
GA+AA/CT+TT	51 (68.9)	6.1	0.33	1.29	1.25 (0.68-2.27)	0.45	0.7
XRCC1 Gln ²⁸⁰ His, G/A & XRCC6 61 C>G	N=95						
GG/CC	15 (15.7)	16.1		1	1		
GA+CG	77 (81.0)	5.8	0.30	1.34	1.51 (0.81-2.8)	0.19	0.57
GA+AA/CG+GG	80 (84.2)	6.2	0.36	1.30	1.48 (0.79-2.74)	0.21	0.7
XRCC1 Gln ²⁸⁰ His, G/A & XRCC7 6721 G>T	N=84						
GG/GG	4 (4.7)	16.1		1	1		
GA/GT	75 (89.2)	8.0	0.5	1.37	1.7 (0.54-5.4)	0.19	0.57
AA/TT	5 (5.9)	21.1	0.81	1.18	-	-	
GA+AA/GT+TT	80 (95.2)	8.0	0.56	1.33	1.78 (0.56-5.63)	0.32	0.7
XRCC1 Gln ²⁸⁰ His, G/A & MUTYH Gln ³²⁴ His, G/C	N=82						
GG/GG	12 (4.7)	6.4		1	1		
GA/GC	69 (89.2)	5.7	0.3	1.38	1.7 (0.54-5.4)	0.19	0.57
AA/CC	1 (5.9)	4.1	0.25	3.0	-	-	
XRCC1 Gln ²⁸⁰ His, G/A & OGG1 Ser ³²⁶ Cys, C/G	N=132						
GG/CC	9 (6.8)	16.1		1	1		

to be cont'd on next page

Table 5.96: Genotypic combinations and overall survival(contd..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R ^a	H.R ^b	p-value	FDR _p
GA/CG	121 (91.6)	6.8	0.02	2.27	2.21 (0.98-5.0)	0.05	0.49
XRCC1 Gln ²⁸⁰ His, G/A & XRCC3 Thr ²⁴¹ Met,C/T	N=85						
GG/CC	10 (11.7)	7.3		1	1		
GA/CT	74 (87.0)	7.2	0.58	0.83	0.82 (0.38-1.74)	0.61	0.89
XRCC6 61 C>G & MUTYH Gln ³²⁴ His, G/C	N=84						
CC/GG	37 (44.0)	9.8		1	1		
CG/GC	47 (55.9)	5.8	0.46	1.18	1.17 (0.69-1.97)	0.5	0.89
XRCC6 61 C>G & OGG1 Ser ³²⁶ Cys, C/G	N=98						
CC/CC	13 (13.2)	15.7		1	1		
CG/CG	84 (85.7)	5.4	0.19	1.51	1.39 (0.71-2.29)	0.32	0.74
XRCC6 61 C>G & XRCC3 Thr ²⁴¹ Met,C/T	N=94						
CC/CC	43 (45.7)	9.0		1	1		
CG/CT	50 (53.1)	6.2	0.59	0.89	0.79 (0.49-1.28)	0.34	0.74
XRCC7 6721 G>T & MUTYH Gln ³²⁴ His, G/C	N=67						
GG/GG	19 (28.3)	10.1		1	1		
GT/GC	47 (70.1)	7.1	0.83	1.04	1.16 (0.62-2.2)	0.63	0.89
XRCC7 6721 G>T & OGG1 Ser ³²⁶ Cys, C/G	N=102						
GG/CC	9 (8.8)	10.2		1	1		
GT/CG	89 (87.2)	6.4	0.57	0.81	0.76 (0.35-1.67)	0.63	0.89
TT/GG	4 (3.9)	16.1	0.17	0.45	0.26 (0.047-1.44)	0.12	0.48
GT+TT/CG+GG	93 (91.1)	7.5	0.51	0.78	0.73 (0.34-1.59)	0.44	0.7
XRCC7 6721 G>T & XRCC3 Thr ²⁴¹ Met, C/T	N=75						
GG/CC	21 (28.0)	10.4		1	1		
GT/CT	54 (72.0)	8.6	0.74	0.91	0.88 (0.49-1.56)	0.04	0.49
MUTYH Gln ³²⁴ His, G/C & XRCC3 Thr ²⁴¹ Met, C/T	N=108						

to be cont'd on next page

Table 5.96: Genotypic combinations and overall survival(contd..)

Genotype	Cases, n(%)	Univariate analysis			Multivariate analysis		
		MST	p-value	H.R ^a	H.R ^b	p-value	FDR _p
GG/CC	53 (49.0)	9.8		1	1		
GC/CT	55 (50.9)	7.2	0.40	1.18	1.36 (0.88-2.12)	0.16	0.57
OGG1 Ser ³²⁶ Cys, C/G & XRCC3 Thr ²⁴¹ Met,C/T	N=91						
CC/CC	11 (12.0)	5.8		1	1		
CG/CT	79 (86.8)	6.2	0.74	1.11	1.13 (0.56-2.28)	0.72	0.89

a-Unadjusted Hazards ratio for Kaplan meier analysis, b-hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG, Histology

5.14.6 High order gene-gene interactions for prognosis of lung cancer patients using survival tree analysis

The analysis is used to find out the high order parametric interactions taking place within the multiple interaction studies using different genes. Further, the effect of these interactions is also analyzed on overall survival of lung cancer patients. In the present study the recursive partitioning method was used to generate a survival tree. The tree obtained by S-tree software shows the involvement of different SNPs in prognosis in various genotypic combinations as shown in fig 5.35 and illustrated in table 5.97. S-tree make use of log-rank algorithm to create the tree. A total of eight terminal nodes were observed in the tree. *OGG1 326* was the most significant factor influencing the survival as it was the root split. The terminal node 3 with shortest MST Of 5.1 months was taken as reference. The tree gives MST for each terminal node and the hazards ration are evaluated respectively further as shown in table 5.97. For Cox-regression analysis hazard ratio was adjusted for various co-factors such as age, gender, smoking, histology, stage, ECOG, KPS. Terminal node 7 with genotypic constitution (*OGG1 326(M)/XRCC3 (M)/XRCC1 194(W)/MUTYH(M)/XRCC1 399(M)/XRCC1 206(M)/XRCC6 (W)*) showed a better prognosis with MST of 7.6 months and a very small hazard rate (H.R=0.073, 95%C.I.=0.0078-0.69, ***p=0.02***). Similarly, terminal node 2 and 4 showed a good prognosis as shown in table 5.97. Figure 5.36 shows the survival curves for various nodes.

Table 5.97: High order gene-gene interactions using survival tree analysis

Node	Genotype in each terminal node	(Dead/ Total) N	MST	Unadjusted H.R. ^a	Log-rank p	Adjusted H.R. ^b	p
Node 1	OGG1 326(W)	22/28	11.1	0.60 (0.29-1.25)	0.12	2.1 (0.82-5.38)	0.12
Node 2	OGG1 326(M)/ XRCC3 (W)	77/86	7.1	0.66 (0.34-1.28)	0.14	0.40 (0.20-0.81)	0.012
Node 3	OGG1 326(M)/ XRCC3 (M)/ XRCC1 194(M)	14/14	5.1	Reference	-	Reference	-
Node 4	OGG1 326(M)/ XRCC3 (M)/ XRCC1 194(W)/ MUTYH(W)	27/36	6.0	0.57 (0.27-1.19)	0.07	0.38 (0.16-0.89)	0.02
Node 5	OGG1 326(M)/ XRCC3 (M)/ XRCC1 194(W)/ MUTYH(M)/ XRCC1 399(W)	8/8	5.7	0.68 (0.28-1.62)	0.35	0.49 (0.11-2.0)	0.34
Node 6	OGG1 326(M)/ XRCC3 (M)/ XRCC1 194(W)/ MUTYH(M)/ XRCC1 399(M)/ XRCC1 206(W)	7/9	5.4	0.58 (0.24-1.38)	0.20	0.88 (0.24-3.1)	0.84
Node 7	OGG1 326(M)/ XRCC3 (M)/ XRCC1 194(W)/ MUTYH(M)/ XRCC1 399(M)/ XRCC1 206(M)/ XRCC6(W)	9/11	7.6	0.68 (0.3-1.54)	0.33	0.073 (0.0078-0.69)	0.02
Node 8	OGG1 326(M)/ XRCC3 (M)/ XRCC1 194(W)/ MUTYH(M)/ XRCC1 399(M)/ XRCC1 206(M)/ XRCC6(M)	11/11	5.2	1.45 (0.63-3.3)	0.3	0.56 (0.19-1.65)	0.29

a- Unadjusted Hazards ratio for Kaplan meier analysis;b-hazards ratio adjusted for age, sex, histology, smoking, stage, KPS, ECOG.

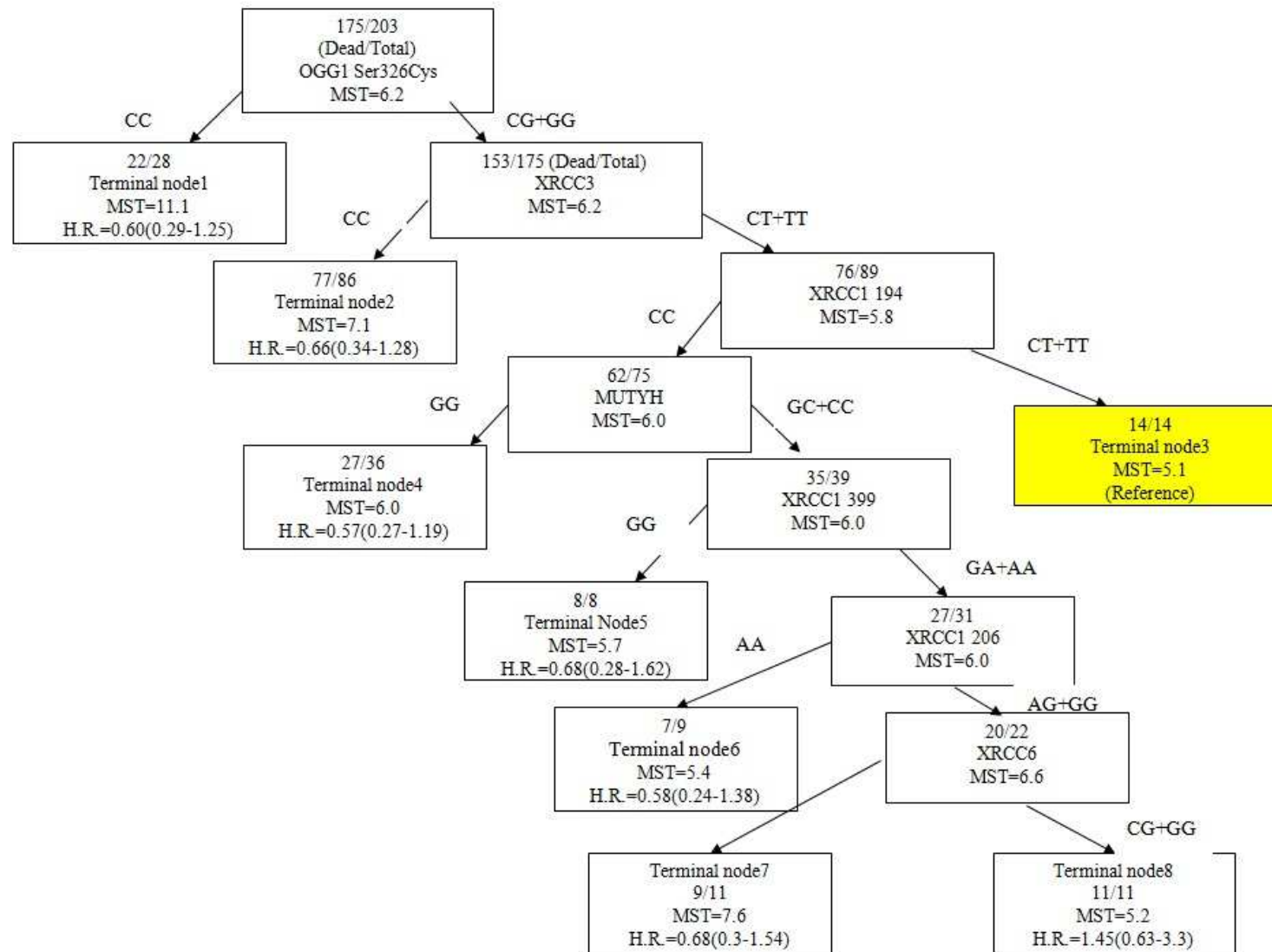


Figure 5.35: CART analysis in polymorphic DNA repair pathway genes. W= homozygous wild type genotype, H+M= heterozygous + homozygous variant genotype

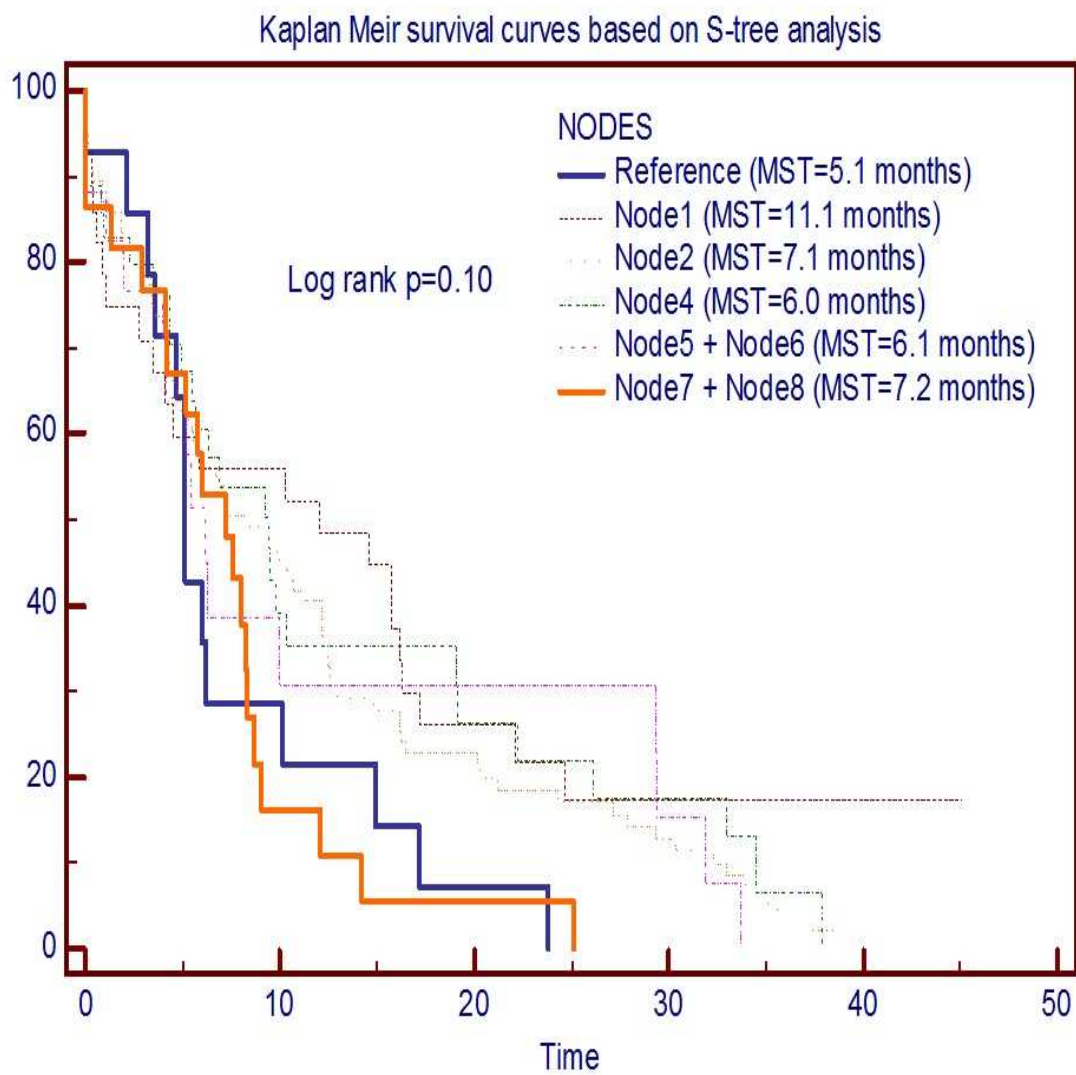


Figure 5.36: Overall survival according to different nodes identified by survival tree analysis

Chapter 6

Discussion

6.1 Discussion

Epidemiological studies with large sample size help us to provide an insight into the relationship between DNA repair genes and the cancer risk. Such studies help us to conclude that irrespective of regulatory pathways and environmental variations, single nucleotide polymorphisms in gene of interest influences the disease. They add to our understanding of public health dimensions due to variations in DNA repair gene.

The integrity of human genome is under constant guards with the help of DNA repair pathways. DNA repair pathways help in the reduction of mutation frequency of cancer related genes. Whenever a deficiency or single nucleotide polymorphism occur in DNA damage repair pathway genes an altered DNA repair gene expression is observed along with suboptimal DNA repair activity (Yuan *et al.*, 2015). This ultimately influences prognosis and survival of lung cancer patients. DNA repair system behave as a double edged sword where impaired DNA repair capacity (DRC) may lead to improper repair capacity to remove DNA lesions created by mutagens and hence might increase the chances of developing cancer, on the contrary same may act beneficial as it would improve the overall survival of cancer patients due to reduced rate of removal of anti-cancerous drugs from DNA, after treatment with platinum agents. Single nucleotide polymorphisms (SNPs) in DNA repair genes force the DRC to vary among individuals. Therefore polymorphic variants can help in the prognosis of lung cancer patients undergoing chemotherapy.

The present study focussed on various genetic variations in BER and DSBR genes which in turn alter the DNA repair efficiency and modulate the susceptibility and survival of lung cancer subjects. Three major BER genes *i.e.* *OGG1*, *MUTYH* and five polymorphic variants of *XRCC1* which are located in different exons are evaluated in association towards lung cancer predisposition. Further, DSBR genes *XRCC3*, *XRCC4*, *XRCC6*, *XRCC7* were also evaluated for their influence on individual's predisposition towards lung cancer. The genetic variants were also examined to evaluate their role in predicting the survival and prognosis of North Indian population. To our knowledge this is the first study

in North Indian population studying these SNPs together in lung cancer predisposition and survival.

6.1.1 BER genes *OGG1* and *MUTYH* and its role in lung cancer predisposition

An increasing body of evidence suggests that oxidative stress is associated with increased initiation and progression of carcinogenesis and ageing (Yan, 2014). DNA repair genes are the candidate susceptibility genes in the etiology of lung cancer, since reactive oxygen species produces a stable guanine product called 8-oxoguanine and repair of mutations involving 8-oxoguanine is a multistep process dependent on proteins of two genes *MUTYH* and *OGG1*. The polymorphic variants *OGG1 Ser³²⁶Cys* (rs1052133) and *MUTYH Gln³²⁴His* (rs3219489) have been extensively studied for their roles towards cancer susceptibility and prognosis. *OGG1 Ser³²⁶Cys* has been observed to have a reduced excision capacity of 8-oxoguanine from double stranded DNA (Nohmi *et al.*, 2005). Our investigative study showed that subjects with single allelic variants (*CG*) for *OGG1* have a strong predisposition towards lung cancer (OR=2.44, ***p*=0.0003**). Our data is also supported by study on lung cancer by (Qian *et al.*, 2011). Studies carried out in colorectal cancer (Lai *et al.*, 2016) and head and neck (Sliwinski *et al.*, 2011) were in concordance with our study. However there are studies which varied from our results where no effect was observed in relation to the heterozygous genotype for *OGG1 Ser³²⁶Cys* (Kohno *et al.*, 2011; Li and Guan, 2011). Study by Marchand *et. al* (Le Marchand *et al.*, 2002) suggested a likely protective effect of heterozygous genotype for Caucasian population and no effect in case of Japanese and Hawaiian population. Study in Brazilian population contradicted our study where a protective effect (OR=0.46) was observed for the heterozygous genotype (Couto *et al.*, 2015). The combined variants exhibited a significant increased risk towards lung cancer (OR=1.99, ***p*=0.0093**) in our study, whereas a study on lung cancer showed no effect for the same genotype (Letkova *et al.*, 2013). Study in North Indian population by Mittal *et.al* in bladder and prostate cancer depicted similar results as our study (Mittal *et al.*, 2012) in relation to combined variant genotype. On the contrary our data suggested that subjects who were having mutant genotype (*GG*) had a protective effect towards lung cancer (OR=0.4, ***p*=0.018**). Our data was similar to a study conducted in turkey population (Karahalil *et al.*, 2008) and a study by De Ruyck *et al.* (De Ruyck *et al.*, 2007). However studies in lung cancer in Chinese population (Mittal *et al.*, 2012) , study by Letkova *et.al.* with emphasis on gender difference (Letkova *et al.*, 2013) and prostate cancer (Mittal *et al.*, 2012) showed varied results from our work. Whereas, some studies contradicted our study where a high risk was observed

for the double allelic variants in lung (Le Marchand *et al.*, 2002), head and neck (Sliwinski *et al.*, 2011) and bladder cancer (Mittal *et al.*, 2012).

MUTYH is responsible for the repair of A-8oxoG. However, germline mutations in *MUTYH* is related to heritable colorectal polyposis (Kantor *et al.*, 2017). However, in our attempt to look for the association of polymorphic *MUTYH* with lung cancer, our study suggested that carriers of single allelic variants (*GC*) possessed an increased risk (OR=2.6, $p<0.0001$). Our data was similar to that reported in lung cancer similar to that in Japanese population (Miyashita *et al.*, 2009). Cancers like colorectal cancer also had similar results where heterozygous genotype was showing an increased risk towards cancer (Przybylowska *et al.*, 2013). Further study by Biyun *et.al* showed no effect of heterozygous genotype towards lung cancer. Similarly study in Glaucoma did not show any association with heterozygotes (Szaflik *et al.*, 2013). The combined variant genotype for *MUTYH* had a high susceptibility towards lung cancer similar to the study in colorectal cancer (Kasahara *et al.*, 2008).

Different, histological subtypes have different clinical manifestations as different histology express different profile of genes. Therefore, it can be concluded that an individuals predisposition towards a certain type of lung cancer might be influenced by different genetic variants. Henceforth, to look for high risk posing subgroup according to histology a stratified analysis was done. Analysis revealed that single allelic carriers in case of *OGG1 Ser³²⁶Cys* exhibited a higher risk towards ADCC subtype of lung cancer and a similar trend was observed for heterozygous genotype for SCLC subtype but with less propensity. Studies carried out in Japanese population indicated no association of *OGG1 Ser³²⁶Cys* with ADCC susceptibility (Kohno *et al.*, 2006; Ito *et al.*, 2002). Similarly a Chinese study was not in concordance with our study where no association was observed for any genotype of *OGG1 Ser³²⁶Cys* for ADCC subtype (Ouyang *et al.*, 2013). A Taiwanese study contradicted our study for all three subtypes of histology. Differing from present work ADCC subtype didn't exhibit any association for single allelic carriers whereas a high risk was observed for subjects with mutant genotype. Taiwanese lung cancer subjects with SQCC histology did not exhibit any risk for *OGG1 Ser³²⁶Cys*, whereas our study exhibited a highly protective effect for SQCC subjects carrying mutant genotype. Furthermore, Taiwanese lung cancer patients with SCLC histology revealed no association for heterozygous genotype and an elevated risk was observed for subjects with mutant genotype (Chang and Hsiao, 2009). Further a study in Turkish population showed no effect towards ADCC subtype whereas a protective effect was observed for double allelic variant genotype towards SCLC subtype similar to our study (Karahalil *et al.*, 2008). A Japanese populace based study for SQCC subjects revealed no association for single allelic

carriers for *OGG1 Ser³²⁶Cys* however a high risk for subjects carrying mutant genotype was observed, which was contradictory to our study (Miyashita *et al.*, 2009).

Histological analysis in case of *MUTYH* exhibited a significant two fold risk for single allelic variants (*GC*) and combined variants (*GC+CC* for ADCC and SCLC subtype whereas three-fold risk was observed for SQCC subtype. These results were in concordance with those obtained in Japanese lung cancer subjects which showed a higher risk for adenocarcinomas while subjects with squamous cell carcinoma exhibited no risk for heterozygous subjects and a high risk for mutant subjects (Miyashita *et al.*, 2009). The carcinogenic exposure, difference in ethnicity, sample size might be the reasons that contribute to the discrepancy between our studies and previous discussed studies. Therefore validation with larger sample size is warranted for every subgroup.

Tobacco smokes accounts for huge generation of ROS species thereby oxidative damage. Studies have shown that patients who are smokers have huge content of 8-oxoGua in their urine (Loft *et al.*, 2012). In this context analysis based on smoking was done for both *OGG1* and *MUTYH* variants. A previous study done in Japanese population reported a high risk for both smokers and non-smokers for polymorphic *OGG1 Ser³²⁶Cys*. Our results similar to the above report did show a high risk for subjects having single allelic carriers in smokers and non-smokers both (Miyashita *et al.*, 2009). Smokers in Turkish population depicted no association while in non-smokers single allelic carriers depicted no association whereas a protective effect was seen for subjects having mutant genotype similar to our study (Karahalil *et al.*, 2008). Chinese study by Li *et al.* revealed that polymorphic *OGG1* did not show any affect on lung cancer irrespective of smoking status (Li and Guan, 2011). A high risk was observed for subjects having single allelic carriers for both smokers and non-smokers. Whereas a protective effect was seen for subjects carrying both allele of variant genotype in smokers and no association for non-smokers in the present study. These results were seen to be in accordance with a study conducted in Japanese lung cancer subjects by Miyashita *et.al* (Miyashita *et al.*, 2009).

The clinicopathological factors were also analyzed in the present study in relation to polymorphic *OGG1* and *MUTYH* genes. A positive correlation was observed between polymorphic *OGG1* and lymph node invasion (***p=0.01***). No association was observed for other clinicopathological factors. For the double combination of *OGG1* and *MUTYH* gene a high risk was observed in the present study for both single allelic carriers and combined variants. On stratifying for histological subgroup a high risk was observed unanimously for all subtypes of lung cancer.

6.1.2 Polymorphism in the gene *XRCC1* and its role towards lung cancer predisposition

XRCC1 is a scaffold protein which provides for recruiting DNA polymerase, ligase and broken ends of DNA. *XRCC1* is reported to be the first identified mammalian gene that protects cells from damaging effects of ionizing radiations (Caldecott, 2008). There are reports that provide an insight into involvement of *XRCC1* in various repair pathways (Caldecott, 2003; Ogi *et al.*, 2010; Gabel *et al.*, 2013). Polymorphisms in structural gene *XRCC1* influence the DNA repair activity of the gene thus the polymorphic *XRCC1* is immensely explored for its role in many diseases including cancer. In the present work polymorphic *XRCC1* is evaluated for its role in lung cancer.

XRCC1 Arg¹⁹⁴ Trp

The present study exhibited a very small variant allele frequency (0.09) in case of *XRCC1 Arg¹⁹⁴ Trp*. This frequency was similar to an earlier study carried out in African-americans where the minor allele frequency was 0.082. Similarly a study in caucasians reported even lower minor allele frequency of 0.059 (David-Beabes and London, 2001). A prior Indian study contradicted our present finding where minor allele frequency was observed to be 0.58 that was found to be similar to Chinese and Korean data (Yoo *et al.*, 2015). Further logistic regression analysis evaluated the odds ratio and it was observed that *XRCC1 Arg¹⁹⁴ Trp* did not impose any significant risk towards lung cancer. This was in concordance with a study on African-Americans where carriers of single allelic variant and combined variant genotype were exhibiting increased risk. On the contrary, subjects with mutant genotype (*TT*) were exhibiting a mild protective effect in Latinos with respect to *XRCC1 Arg¹⁹⁴ Trp* polymorphism (Chang *et al.*, 2008). Caucasian studies in Southern Italian population and German population were also in concordance with our study where no effect was observed towards lung cancer (Improta *et al.*, 2008; Schneider *et al.*, 2005). Similar results were observed in other cancers like gastric (Lee *et al.*, 2001), breast (Deligezer and Dalay, 2004), head and neck (Kumar *et al.*, 2012), prostate and bladder cancer (Mittal *et al.*, 2012). Meta-analytical studies also strongly supported our conclusions where *XRCC1 Arg¹⁹⁴ Trp* was not observed to be correlated towards lung cancer risk (Zhang *et al.*, 2014; Han Su, 2016). *XRCC1 Arg¹⁹⁴ Trp* is a non-synonymous mutation found in possible binding domain of interacting partners. Change in amino acids would represent a transition from positively charged amino acid to hydrophobic amino acid which would affect binding and ultimately DNA repair capacity (Ladiges, 2006). However, there are studies proposing that the mutant genotype for *XRCC1 Arg¹⁹⁴ Trp* possessed an increased risk towards lung cancer as observed in Indian (Pachouri *et al.*,

2007), Chinese (SHEN *et al.*, 2005) and in East Chinese Han population (De Ruyck *et al.*, 2007; Du *et al.*, 2016). Chinese study on lung cancer reported *XRCC1 Arg¹⁹⁴ Trp* could be used as molecular marker in case of lung cancer susceptibility (Zhu *et al.*, 2015). In cancers other than lung such as breast (Smith *et al.*, 2003), it was conveyed that mutant genotype (*TT*) has an increased risk. Two Chinese (Du *et al.*, 2016; Han *et al.*, 2015) studies conveyed an overall high risk of NSCLC in case of *XRCC1 Arg¹⁹⁴ Trp*.

XRCC1 Pro²⁰⁶ Pro

Synonymous mutations do not change amino acids therefore bring no impact on protein function, *XRCC1 Pro²⁰⁶ Pro* is a synonymous substitution and there are not much reports on its relevance towards lung cancer. One earlier study in Chinese population has discussed about the allelic (A=0.90, G=0.10) frequency distribution which was found to vary from ours results (A=0.70, G=0.30) (Yin *et al.*, 2007). A European study evaluated the involvement of *XRCC1 Pro²⁰⁶ Pro* in lung, bladder cancer and leukemia. Their results were consistent with our analysis where heterozygotes imposed a very moderate risk towards lung cancer and double allelic variants showed no effect at all. In case of bladder cancer no effect was observed however, in case of leukemia moderate protective effect was observed (Matullo *et al.*, 2005). On the contrary, a Chinese population based study revealed that carriers of minor allele variant possessed an increased risk towards lung cancer (Yin and Vogel, 2007). Study conducted in non-smoking Chinese women, Pakistani subject with head and neck cancer imitated the same pattern (Yin *et al.*, 2009; Mahjabeen *et al.*, 2013) as discussed above. Studies in breast cancer (Yin and Wang, 2012), basal cell carcinoma (Yin *et al.*, 2003) and bladder cancer (Sak *et al.*, 2007) showed no association of *XRCC1 Pro²⁰⁶ Pro* towards risk for cancer susceptibility.

XRCC1 Arg²⁸⁰ His

XRCC1 Arg²⁸⁰ His is not much studied polymorphism compared to *XRCC1 Arg¹⁹⁴ Trp* and *XRCC1 Arg³⁹⁹ Gln*. Study by Takanami *et al.* suggested that presence of histidine allele of *XRCC1 Arg²⁸⁰ His* forced a compromised DNA repair activity of *XRCC1* in a MMS sensitive CHO cell line (Takanami *et al.*, 2005). Study conducted in lung cancer patients supported that the variant allele genotype is associated with increased lung cancer risk (Ratnasinghe *et al.*, 2001). Studies in other cancers, like gastric in Korean (Jin *et al.*, 2015), hepato-cellular carcinoma in Indian population (Kiran *et al.*, 2009) were consistent with our study where variant allele (may be single allelic variant or double allelic variants) imposed an increased risk towards cancer. However, there are reports which have observed the opposite in two Chinese studies (Yin and Vogel, 2007), (Yin *et al.*, 2009) where a protective effect was observed for the mutant genotype. Similarly, a Hispanic study revealed a protective effect for carriers of mutant genotype (*AA*) and no

effect for heterozygotes or combined variants towards lung cancer (Chang *et al.*, 2008). Study conducted by Ruyck *et al.* observed that subjects carrying one variant allele showed a protective effect towards lung cancer (De Ruyck *et al.*, 2007). Further, a protective effect was observed in prostate (Xu *et al.*, 2007), for mutant genotype and no effect in bladder (Stern *et al.*, 2001) and nasopharyngeal cancer (Cho *et al.*, 2003). Meta-analytical data on *XRCC1 Arg²⁸⁰His* revealed that subjects with mutant genotype showed an increased risk towards lung cancer whereas individuals with single allelic variant or combined variant genotypes had no effect towards lung cancer susceptibility (Guo and Yang, 2013). Our results were in line with the above meta-analytical study.

XRCC1 Arg³⁹⁹Gln

The polymorphic variant *XRCC1 Arg³⁹⁹Gln* is located in the C-terminal domain of the gene. Functional studies revealed no difference in ability to repair single strand breaks created by alkylating agent in CHO cell lines transfected with either Arg allele or Gln allele of *XRCC1 Arg³⁹⁹Gln* (Taylor *et al.*, 2002). Further high chromosomal deletion was observed in subjects exposed to UV-light (Au *et al.*, 2003). In this study, we report that subjects heterozygous for the *XRCC1 Arg³⁹⁹Gln* polymorphism showed no effect towards lung cancer, whereas subjects homozygous for the variant A allele (AA genotype) depicted a protective effect with marginal significance ($p=0.06$). Our results were in agreement with studies done in Caucasians and African-Americans (David-Beabes and London, 2001), Chinese (SHEN *et al.*, 2005), and subject from Northern Spain (López-Cima *et al.*, 2007). Recent meta-analytical report also suggested the involvement of *XRCC1 Arg³⁹⁹Gln* in individuals susceptibility in Caucasian subjects (Chen *et al.*, 2015). On the same lines studies conducted in Turkish and North Eastern Chinese population partially supported our results where no effect was observed for genotypic combinations of *XRCC1 Arg³⁹⁹Gln* towards lung cancer risk (Yin and Vogel, 2007; Karkucak *et al.*, 2012). Meta-analytical studies conducted have also shown consistency with our observations where protective effect (Wang and Yang, 2009) or no association (Kiyohara *et al.*, 2006; Hung and Hall, 2005) was observed for *XRCC1 Arg³⁹⁹Gln* towards lung cancer risk. *XRCC1 Arg³⁹⁹Gln* has been exhaustively studied but the results have been inconclusive and vary from population to population. A Korean study revealed that individuals with double variant genotype (AA) for *XRCC1 Arg³⁹⁹Gln* possessed a high risk towards lung cancer (Park *et al.*, 2002). A recent meta-analytical study confirmed that the risk of lung cancer due to mutant genotype of *XRCC1 Arg³⁹⁹Gln* strengthens according to ethnic-stratification for Asians (Wang and Cai, 2014). This paradigm is supported by a Chinese (Li *et al.*, 2013) and Indian studies (Natukula *et al.*, 2013; Uppal *et al.*, 2014).

XRCC1 Gln⁶³²Gln

In line with above results another synonymous mutation taken into consideration was *XRCC1 Gln⁶³²Gln*. There was a significant difference in genotypic distribution between cases and controls which was in line with study of Mahjabeen *et.al* (Mahjabeen *et al.*, 2013) where an increased risk towards lung cancer was reported. Similarly, study on non-smoker Chinese women revealed an increased risk towards lung cancer for individuals carrying at-least one variant allele (Yin *et al.*, 2009). A Caucasian study implied a trend towards increased risk in African-Americans and a protective effect in Latinos (Chang, 2008) which was partially trending towards our study. A study conducted in bladder cancer also supported our observation where a trend of increased risk was observed. Due to lack of data present there is very less known about this polymorphism and its role in modulating lung cancer. Although the molecular mechanism for both the synonymous mutations of *XRCC1* gene i.e. *Pro²⁰⁶Pro* and *Gln⁶³²Gln* has so far not been deciphered. Both of these silent polymorphisms does not result in an amino acid substitution and hence the possible mechanism for these synonymous mutations to modulate lung cancer susceptibility might be because of change in translation efficiency or m-RNA expression (Yin *et al.*, 2003). On the other hand, non-synonymous mutations like *XRCC1 Arg³⁹⁹Gln* is found in PARP binding domain whereas, *XRCC1 Arg¹⁹⁴Trp* and *XRCC1 Arg²⁸⁰His* are found in linker region, affecting complex assembly or repair efficiency (London, 2015).

Polymorphic *XRCC1* and its relation to lung cancer predisposition on the basis of histological subtypes

Furthermore in the present study, lung cancer risk was evaluated in various subgroups based on histological subdivision. Each of these histological subtype has different etiology and pathogenesis. They are observed to have different alteration in expression of various genes therefore, there always occur a difference in individuals' predisposition towards certain type of lung cancer. In case of *XRCC1 Arg¹⁹⁴Trp* in the present study increased risk towards ADCC for subjects carrying at-least one copy of the variant genotype was observed. These results were in agreement with a study in North Indian population (Pachouri *et al.*, 2007). There are reports implying a high risk of mutant genotype towards ADCC (Du *et al.*, 2016; Cătană *et al.*, 2015). A protective or no effect was observed for SQCC and SCLC subjects in a Caucasian study (Schneider *et al.*, 2005) and European study (Hung and Brennan, 2005). These inconsistencies in the above results, might be due to ethnic differences or small sample size. Subjects with polymorphic variant *XRCC1 Pro²⁰⁶Pro* did not show any association with lung cancer in case of subjects with ADCC and SQCC subtype. However, study subjects with SCLC subtype and carrying combined variant genotype displayed a high risk towards lung cancer. Study on histological stratification revealed that *XRCC1 Arg²⁸⁰His* had no risk towards ADCC

or SQCC subtype of lung cancer (Hung and Brennan, 2005). Our study was in agreement with Schinder *et.al* (Schneider *et al.*, 2005) where double allelic variants were showing an increased risk in ADCC and SQCC subtype of lung cancer.

A large scale study in six eastern European countries along with a study in Chinese Han population found no effect towards either ADCC or SQCC subtype of lung cancer for both single allele or double allelic variants of *XRCC1 Arg³⁹⁹Gln* (Hung and Brennan, 2005; Zhang and Miao, 2005). These previous two studies were partially in concordance with our study where subjects with heterozygous genotype for *XRCC1 Arg³⁹⁹Gln* for ADCC subjects in our study showed a marginal protective effect ($p=0.08$) and did not show any association for SQCC subtype. Study on lung cancer in Chinese population showed similar risk for mutant subjects with ADCC subtype as in our study indicating a trend towards mild protective effect. Further, our study revealed a protective effect towards SQCC subtype ($p=0.05$) for double allelic variants *i.e.* mutant genotype. However, some studies contradicted our findings, where an increased risk was observed towards SQCC subtype in Korean population (Park *et al.*, 2002). Study in Hispanic and non- Hispanic population strongly contradicted our study where an increased risk was observed towards ADCC subtype of lung cancer (Divine *et al.*, 2001).

Polymorphic *XRCC1* and its relation with lung cancer on the basis of smoking

Smoking is considered a major role player in lung cancer, this factor may interact with other genetic variabilities differently to modulate the individual's predisposition towards the oncogenesis of lung tissue. Therefore in order to evaluate the role of smoking we analyzed the risk exhibited by smokers and non-smokers for different genotypic combinations. In the present study smokers in case of polymorphic variants of *XRCC1* showed no association with lung cancer except for *XRCC1 Gln⁶³²Gln*. In case of three genetic variants namely *XRCC1 Pro²⁰⁶Pro*, *XRCC1 Arg²⁸⁰His*, *XRCC1 Arg³⁹⁹Gln* non-smoker subjects showed a protective effect. Whereas, non-smokers for *XRCC1 Arg¹⁹⁴Trp* and *XRCC1 Gln⁶³²Gln* were found to be at high risk of developing lung cancer. A previous study reported similar results for the three polymorphic variants *XRCC1 Arg¹⁹⁴Trp*, *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg²⁸⁰His* where no association was observed for smokers. In case of non-smokers subjects with polymorphic variant of *XRCC1 Arg¹⁹⁴Trp* acquired a high risk towards lung cancer which was in concordance with our study. Polymorphic variant of *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg²⁸⁰His* contradictory to our study exhibited no effect (Hung and Brennan, 2005). Chinese study for polymorphic variant *XRCC1 Arg¹⁹⁴Trp* depicted a protective effect for never smokers and a high risk for smokers, which was not in accordance with our study (Han *et al.*, 2014). Studies in Indian population

for polymorphic variant *XRCC1 Arg³⁹⁹Gln* showed a positive association of smoking with the polymorphism (Natukula *et al.*, 2013) which was contradictory to our result where no association was observed. Another Indian study depicted a positive correlation with smoking (Uppal *et al.*, 2014) . The present study showed a very high risk for the polymorphic variant *XRCC1 Gln⁶³²Gln* for both smokers and non-smokers.

Combinatorial genotypic analysis was carried out to find out the higher risk subgroup with a particular genotypic combination. Therefore, we have also analyzed possible interactions between SNPs in the *XRCC1* gene. Out of all the combinations analyzed only certain groups were modulating the lung cancer risk. Our study revealed that double combinations of *XRCC1 Gln⁶³²Gln* with *XRCC1 Pro²⁰⁶Pro* significantly elevated lung cancer risk ($p < 0.00041$). Whereas, other combinations were either elevating the risk or having no effect towards lung cancer without any significance. There are very few studies regarding the effect of combinations of SNP's of *XRCC1* towards lung cancer susceptibility, however one of the study in hepatocellular carcinoma (HCC) illustrated that subjects carrying single allelic variants for the following combinations: *XRCC1 Arg¹⁹⁴Trp* and *XRCC1 Arg²⁸⁰His*, *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg¹⁹⁴Trp*, *XRCC1 Arg²⁸⁰His* and *XRCC1 Arg³⁹⁹Gln*, showed an increased susceptibility towards HCC (Kiran *et al.*, 2009). Combinatorial assessment of *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg¹⁹⁴Trp* in glioma showed that individuals with mutant genotype for the combination imposed an increased risk towards glioma (Hu *et al.*, 2011). However, on the contrary Indian data in head and neck cancer revealed a highly protective effect for the the genetic combination of *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg¹⁹⁴Trp* for both heterozygous and mutant genotype (Kumar *et al.*, 2012). Our results have shown no effect of the mutant or variant genotype for the combination of *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg¹⁹⁴Trp* towards lung cancer susceptibility. However, for the double combination of *XRCC1 Arg³⁹⁹Gln* either with *XRCC1 Pro²⁰⁶Pro* or *XRCC1 Gln⁶³²Gln* respectively, a 2-fold risk was observed in case of single allelic variants. We also evaluated the effect of triple combinations on lung cancer and analyzed that the carriers of heterozygous genotype for the combination of *XRCC1 Arg¹⁹⁴Trp*, *XRCC1 Pro²⁰⁶Pro* and *XRCC1 Gln⁶³²Gln* imposed a high risk (OR=6.12, $p = 0.0002$).

Haplotype analysis for *XRCC1*

Furthermore, haplotype analysis for *XRCC1* variants was carried out. There are very few reports scrutinizing the haplotype analysis of five SNPs together. This is the first report giving us an insight into the five SNPs together in Indian data. Here we perceived, haplotype blocks carrying minor allele for synonymous SNPs imposed a higher risk in comparison to blocks carrying minor allele for non-synonymous SNPs. In line with this

haplotype block 11 (CGAGG) was showing six-fold increased risk towards lung cancer (OR=6.11, $p=1.48e-010$) which carried all wild alleles except for *XRCC1 Pro*²⁰⁶*Pro* and *XRCC1 Arg*²⁸⁰*His*. Likewise, haplotype block 12 (CGGGG) carrying mutant allele only for *XRCC1 Pro*²⁰⁶*Pro* showed a higher risk (OR=3.054, $p=0.000104$) in comparison to other haplotype blocks. This depicted that synonymous polymorphism is important in disease susceptibility. Similarly, haplotype blocks carrying minor allele for *XRCC1 Gln*⁶³²*Gln* e.g. haplotype block 2 (CAGAA) (OR=3.6, $p=1.18e-008$) and haplotype block 1 (CAAAA) (OR=2.68, $p=2.06e-006$) showed an increased risk in comparison to other blocks. Therefore, these observations helped us to deduce that haplotypes carrying minor allele for synonymous variants (*XRCC1 Pro*²⁰⁶*Pro*, *XRCC1 Gln*⁶³²*Gln*) were essentially at higher risk compared to other haplotype blocks as discussed in two Chinese studies (Yin and Vogel, 2007; Yin *et al.*, 2009). The current study also evaluated relationship of these five SNPs with clinicopathological parameters and out of all the genetic variants association was seen in case of subjects with double allelic variant genotype of *XRCC1 Arg*²⁸⁰*His* and lymph node involvement ($p=0.012$). Thus *XRCC1 Arg*²⁸⁰*His* variant might be playing a role in tumor progression if not in modulating the susceptibility towards lung cancer.

MDR and CART analysis for *XRCC1*

The non-parametric methods such as MDR and CART were used to evaluate the cumulative role of polymorphic *XRCC1* variants in modulating lung cancer susceptibility. Consistent with our regression results, we found that in MDR analysis *XRCC1 Gln*⁶³²*Gln* was the highest risk factor with a CVC of 10/10 and prediction error of 0.41 along with $p=0.07$. *XRCC1 Gln*⁶³²*Gln*, *XRCC1 Pro*²⁰⁶*Pro* and *XRCC1 Arg*²⁸⁰*His* in combination act as the best three factor risk model with 10/10 CVC and prediction error of 0.34 and $p=0.01$. Similarly, the data mining statistics using CART also used *XRCC1 Gln*⁶³²*Gln* as first split and next split was on *XRCC1 Pro*²⁰⁶*Pro* and *XRCC1 Arg*²⁸⁰*His* which was in parallel with MDR results. *XRCC1 Gln*⁶³²*Gln* came out to be the most significant risk posing factor. CART analysis divides the risk groups in terminal nodes and calculate the risk for every node which help us to evaluate the risk factors as in a study on NER genes (Chen *et al.*, 2007). Node1 exhibited the highest risk of 29 folds with a combination of *XRCC1 Arg*²⁸⁰*His* W and (*XRCC1 Gln*⁶³²*Gln*M).

6.1.3 DSBRR genes (*XRCC4*, *XRCC6* and *XRCC7*) polymorphism and lung cancer predisposition

DSBR gene *XRCC6* and *XRCC7* in lung cancer predisposition

The NHEJ genes *XRCC6* and *XRCC7* helps in keeping the preservation of genomic in-

tegrity which is essential to prevent initiation and development of tumor. Variations or single nucleotide polymorphism in DNA repair genes may force them to escape cell cycle surveillance and lead to suboptimal DNA repair and ultimately carcinogenesis. To the best of our knowledge this is the first study in North Indian subjects studying the effect of two polymorphic genes on lung cancer susceptibility. Meta-analytical study revealed that there was immense difference in the minor allele frequency (G) on the basis of ethnicities. Our study population had MAF of 0.29 which was similar to as observed in both Asian population (0.05) and European subjects (0.39) (Jia *et al.*, 2015). Further, in the current study for $XRCC6$ $61C>G$ a trend of mutant genotype (GG) towards susceptibility of lung cancer (OR=1.97, $p=0.08$) was observed. Studies on hepatocellular carcinoma (Li and Yang, 2011), renal cell carcinoma (Chang *et al.*, 2012), acute myeloid leukemia (Wang and Wang, 2009), gastric (Yang *et al.*, 2011) and breast (He *et al.*, 2012) cancer were in line with our results where double allelic variant genotype (GG) was found to be associated with high risk for the respective cancers. On the contrary studies done in lung (Tseng *et al.*, 2009) and esophageal cancer (Li and Suo, 2012) depicted no association of cancer susceptibility with $XRCC6$ $61C>G$ polymorphism. Lung cancer specific study conducted in Taiwan based population showed no association of SNP with lung cancer (Hsia *et al.*, 2012). $XRCC6$ act as scaffold to recruit all NHEJ factors to the damaged site. $XRCC6$ $61C>G$ is a functional polymorphism present in the coding region of promoter (Bau *et al.*, 2011). Impairments in $XRCC6$ could lead to improper DNA repair which would ultimately promote apoptosis, chromosomal rearrangements, genomic instability further giving way to carcinogenesis (Beggs *et al.*, 2012). Mutant genotype of polymorphic $XRCC6$ $61C>G$ showed involvement in male infertility (Jahantigh and Hosseinzadeh Colagar, 2017). Metanalysis conducted in relation to $XRCC6$ $61C>G$ suggested that the polymorphism is associated with increased risk in Asian population (Jia *et al.*, 2015).

The next polymorphism in the other upstream genes studied is $XRCC7$ $6721G>T$. $XRCC7$ is also known as catalytic subunit of DNA-PK. There are reports that describe deficiency in DNA-PK as responsible for many clinical conditions. However this polymorphism was also found to be implicated towards prostate (Mandal *et al.*, 2010), bladder (Wang *et al.*, 2008; Zhi *et al.*, 2012) glioma (Wang *et al.*, 2004) and hepatocellular carcinoma (Hsieh *et al.*, 2015). Whereas, our study revealed $XRCC7$ $6721G>T$ is not associated with lung cancer risk. Study carried out in Chinese population on lung cancer validated our study where similar to our study both single allelic variant (GT) (OR=0.76) and double allelic variants (TT) (OR=1.33, $p=0.46$) did not pose any significant risk towards lung cancer (Hu *et al.*, 2008). The findings in the present work are in concordance with studies in other cancers such as breast (Nasiri *et al.*, 2012) and renal cancer (Hirata

et al., 2006). On the contrary, Hsia *et al.* found a protective effect of *XRCC7 6721G>T* towards lung cancer (Hsia *et al.*, 2014).

Polymorphic *XRCC6* and *XRCC7* towards lung cancer in relation to histology and smoking status

Due to variation in the pathogenesis of different histological subtypes of lung cancer, they exhibit slightly different clinical manifestations; also they are mediated through distinct alterations in expression of various genes and proteins. In the present study variant genotype for *XRCC6 61C>G* did show a high risk towards ADCC subtype. The risk for variant genotype was observed to increase further for SQCC and SCLC subsequently. *XRCC7 6721G>T* also did show a high risk for SCLC subtype. Further, due to less sample size, in the stratified histologies the findings in the present study need further validation. In order to evaluate the role of smoking which is considered a major role player of lung cancer, we analysed the prevalence of the variant genotypes in smokers and non-smokers, as it is important in detecting the gene-environment interactions in order to categorize the subgroups who are at higher risk of acquiring lung cancer due to environmental factors. In the current analysis smokers with variant genotype contributed towards lung cancer risk whereas non-smokers with hetero-genotype conferred a protective effect towards lung cancer. Similar to *XRCC6* smokers in case of polymorphic *XRCC7 6721G>T* were found to be slightly high risk whereas no affect was observed in non-smokers. On the same lines a significant difference was observed in the genotypic frequencies of cases and controls if they had a variant genotype in Taiwan male smokers whereas non-smokers had no difference indicating interaction with smoking (Hsia *et al.*, 2014). The present study also evaluated the relationship of these SNPs' with clinico-pathological parameters and out of the two genetic variants a positive correlation was observed in case of subjects with combined variant genotype of *XRCC6 61C>G* and clinical staging (OR=1.49, p=0.09). Combinatorial genotypic analysis was carried out to find out the high risk subgroups with a particular genotypic combinations. None of the genotypic group for the combinations revealed any risk towards lung cancer. Further, analysis for histological subtype for the combination revealed that subjects diagnosed with SCLC subtype possessed a high risk towards lung cancer.

DSBR gene *XRCC4* and its association for lung cancer predisposition

DSBR gene *XRCC4* is a NHEJ protein involved in VDJ recombination. *XRCC4* deficient mice resulted in pleiotropy which resulted in neuronal apoptosis, embryonic lethality and impaired cellular proliferation (Gao *et al.*, 2000). In the present study, polymorphic *XRCC4* did not show any association with lung cancer. Many studies have reported the same in many cancers including lung as stated ahead. A Taiwan populace based study

also gave the same insight about this polymorphic site where no significant difference was observed between the genotypic frequencies of cases and controls for lung cancer subjects (Hsu *et al.*, 2009). Similar results were observed in breast (Saadat and Saadat, 2015), bladder (Chang and Chang, 2009), and gastric cancer (Chiu *et al.*, 2008) where no difference in the distribution of genotype was observed. In contrast to the above studies, an Iranian population suggested a high frequency of *Del* allele of the polymorphic variant, and the prevalence was much higher than in the Asian population (Fallahzadeh-Abarghoeei *et al.*, 2015). Study in prostate cancer in Indian population suggested a high risk due to both deletion alleles (Mandal *et al.*, 2011). Similarly, Taiwan population in oral cancer depicted an increased risk due to polymorphic *XRCC4* (Chiu *et al.*, 2008). Significant difference between UBC and controls was also observed in North Indians (Mittal and Gangwar, 2012).

6.1.4 DNA repair genes and their role in overall survival and objective response of lung cancer patients receiving platinum based chemotherapy

Single nucleotide polymorphisms (SNPs) in genes of DNA repair pathways may alter the expression of genes and therefore alter the DNA repair activity. This in turn influences the prognosis, effectiveness of chemotherapeutic treatment and ultimately survival of patients. Polymorphic *XRCC1* is observed to be critical in repair of damages induced by platinum drugs such as cisplatin or carboplatin, as experimentally proven by finding transcripts abundance in response to cisplatin resistance in NSCLC cell lines (Mohrenweiser *et al.*, 2002). Mice subjects with polymorphic *XRCC4* showed a pleiotropic expression (Gao *et al.*, 2000). However, the inferences about the prognostic implications of BER and DSBRR pathway genes vary for ethnic differences. The present work provides a data of various DNA repair genes and their implications in overall survival, survival according to stratified histology, chemotherapy, and objective response for subjects receiving platinum based doublet chemotherapy, in North Indian population.

OGG1 and MUTYH: Findings, from the current study suggested a lack of any significant association of the variants of BER or DSBRR genes with overall survival of lung cancer patients. It was observed that better survival was observed in case of *OGG1 Ser³²⁶Cys* for mutant genotype (*GG*) compared to wild type (*CC*) (16.1 vs 12.0). However the hazard rate did not show much alteration. A Chinese study reported contradictory results to our study where patients with G allele were specifically associated with poor survival (Su *et al.*, 2015). The results were strengthened specifically in female, early stage and light

smokers (Su *et al.*, 2015). On the same lines mutant type (*CC*) for *MUTYH Gln³²⁴His* proposed a better survival when compared to wild type (*GG*) genotype (15.3 vs 7.6) whereas the death rate was not affected. Survival for lung cancer is not much evaluated in populations in relevance to the two BER genes *i.e.* *OGG1* and *MUTYH*. However, one of the recent studies clearly indicated that *OGG1 Ser³²⁶Cys* was not associated with overall survival however, *OGG1 Ser³²⁶Cys* was associated with shorter PFS. Further the same study detailed that *OGG1 Ser³²⁶Cys* along with *XRCC1 Arg³⁹⁹Gln* showed an association with overall survival (Peng *et al.*, 2014). Furthermore, *OGG1* is studied to have a prognostic role in stratifying acute myeloid leukemia patients (AML) those who are likely not to respond to chemotherapy but to novel therapeutic approaches (Liddiard *et al.*, 2010). Study on rectal cancer has presented that *OGG1 Ser³²⁶Cys* acts as an important factor towards tumor response in patients treated with chemotherapy as neo-adjuvant therapy (Cecchin *et al.*, 2011). However it was observed not to be associated with radiotherapy (Monteiro *et al.*, 2005).

Furthermore, *MUTYH* is immensely studied in relation to colorectal cancer as alteration in both the alleles of *MUTYH* may lead to the development of colorectal cancer (Kashfi *et al.*, 2013). Its been studied that patients with gastric cancer exhibiting low *MUTYH* expression showed a poor outcome when compared to the ones expressing high expression of *MUTYH*. Further reduced expression of *MUTYH* acted as an independent predictor of poor survival in gastric cancer patients (Shinmura *et al.*, 2011). In case of polymorphic variant *OGG1Ser³²⁶Cys* a trend was observed for SCLC subtype towards an increased death rate for lung cancer subjects. Contrary to our study variant allele was associated with poor prognosis in Adenocarcinomas. In case of *MUTYH Gln³²⁴His* for the three histology MST was mostly higher for mutant type when compared to wild and heterozygous genotype as detailed in the results. Prior study on *MUTYH* in lung cancer showed no association with lung cancer (Su *et al.*, 2015).

Influence of genetic variants of *XRCC1* on survival

For the five polymorphic variants of *XRCC1* in our study it was observed that patients with mutant genotype (*AA*) for *XRCC1Arg³⁹⁹Gln* have an increased OS period (MST=9.6 months) in comparison to carriers of wild (*GG*) type genotype or heterozygous (*GA*) genotype. Studies conducted in Chinese population (Zhang and Ma, 2014) and in Taiwan population (Liao *et al.*, 2012) supported our conclusion where both studies have reported an increased overall survival for mutant genotypes of *XRCC1Arg³⁹⁹Gln*. Another Chinese report supported our conclusion where mutant subjects had an increased survival (Liu *et al.*, 2014). Similar results were also obtained in Japanese study (Osawa *et al.*, 2012). Studies in other cancers like hepatocellular carcinoma (Wu *et al.*, 2014), ovarian

cancer (Li and Li, 2013), are also in agreement with our study. The predictive role of *XRCC1 399* is considered controversial as there are studies where mutant genotype (AA) is associated with shorter overall survival. One of the initial study by Gurubhagavatula *et.al* (Gurubhagavatula *et al.*, 2004) represented shorter survival of patients with mutant genotype (AA) indicating poor prognosis towards lung cancer. Studies done in Chinese population (Jin *et al.*, 2014), Indian population (Sreeja *et al.*, 2008) also supported the above paradigm. Investigations in other cancers like hepatocellular carcinoma (Li *et al.*, 2012), bladder cancer (Sacerdote *et al.*, 2013), colon cancer (Zaanan *et al.*, 2014) also revealed a shorter survival of patients with mutant genotype (AA). A study conducted in head and neck cancer (Ang *et al.*, 2011) suggested an increased expression of *XRCC1* is associated with poorer survival. There are studies which report no effect of *XRCC1 399* on overall survival of lung cancer subjects undergoing chemotherapy (Yuan *et al.*, 2010; Yao *et al.*, 2009). However, along with the studies discussed above, meta-analytical data suggest a predictive role of *XRCC1 399* on overall survival of lung cancer patient (Cao *et al.*, 2014). Further, on analyzing pathological features for their association with overall survival our study revealed subjects carrying double allelic variants (AA) had longer MST (9.4) months for ADCC subtype. These results were in concordance with study done in non-smoking female adenocarcinoma patients (Yin *et al.*, 2009).

The current data from our study showed no association of *XRCC1 Arg¹⁹⁴ Trp* with death rate or survival similar to as revealed by other studies (Yuan *et al.*, 2010; Zhang and Ma, 2014; Gurubhagavatula *et al.*, 2004). Whereas, there are reports that predict the association of mutant genotype (TT) of *XRCC1 Arg¹⁹⁴ Trp* with better survival (Zhao *et al.*, 2013; Zhao and Chen, 2015). Cancers like ovarian (Li and Li, 2013), hepatocellular carcinoma (Wu *et al.*, 2014), and bladder cancer (Sacerdote *et al.*, 2013) also depicted an increased survival for mutant genotype. Furthermore, histological stratification also revealed no association of *XRCC1 Arg¹⁹⁴ Trp* with any histology. The *XRCC1 Arg¹⁹⁴ Trp* amino acid substitution resides in linker region separating DNA polymerase B domain and PARP interacting domain. The SNP may be responsible for disrupting the interaction and effect the functionality, however the reason for observing no effect could be the source of origin, frequency of polymorphic variant in population as we had very few mutant genotype (TT) in our study group.

There is plethora of reports on association of non-synonymous mutations with survival or sensitivity towards chemotherapeutic drugs for cancer patients. To our complete knowledge there is no study revealing any association between synonymous mutations and survival. In our study we have made an attempt to analyze *XRCC1Pro²⁰⁶Pro*, and *Gln⁶³²Gln* for their any involvement in survival. On investigating *XRCC1Pro²⁰⁶Pro* it

was observed that though the MST was high for mutant genotype (*GG*) (MST=8.8) the hazards ratio did not show any association with overall survival. Furthermore, in case of *XRCC1 Gln⁶³²Gln* no association was observed with overall survival for any genotypic combination. *XRCC1Arg²⁸⁰His* is evaluated for its implication in survival. Our observations suggested no influence on overall survival. This investigation of our study is supported by other studies in lung cancer (Yuan *et al.*, 2010; Jin *et al.*, 2014). An ovarian cancer study also supported our results where no association was observed with death risk (Li and Li, 2013). However some contradictory results were observed in studies where mutant genotype (*AA*) were having a prolonged survival (Yoon *et al.*, 2005; Zhao and Chen, 2015). On the contrary studies in ovarian cancer (Cheng *et al.*, 2012) gastric cancer (Hu *et al.*, 2016) proposed that mutant genotype (*AA*) possessed an increased risk towards death.

Polymorphic *XRCC3*, *XRCC4*, *XRCC6* and *XRCC7* and its role in survival of lung cancer patients

Variant of DSBRR gene *XRCC3* has been observed to be associated with higher DNA adducts level irrespective of smokers or non-smokers. In respect with this, in the present study we have assessed the role of *XRCC3 Thr²⁴¹Met* in lung cancer prognosis. The polymorphism *XRCC3 Thr²⁴¹Met* has been studied before, for its role in survival (De las Penas *et al.*, 2006; Zhou *et al.*, 2010; Shen *et al.*, 2013). Our results did not show significant association with overall survival similar to the study conducted in Japanese and Chinese population (Osawa *et al.*, 2012; Jin *et al.*, 2014). The functional significance of the SNP is hypothesized to have a changed phosphorylation site due to amino acid change (Bewick *et al.*, 2006). Similarly study by Ke *et al.* did not show any association of *XRCC3 Thr²⁴¹Met* with lung cancer (Ke *et al.*, 2012). *XRCC3 Thr²⁴¹Met* was observed to affect survival of patients in breast and head and neck cancer (Bewick *et al.*, 2006; Farnebo *et al.*, 2015). A reduced survival was observed for pancreatic cancer patients in association with variant alleles of *XRCC3* (Li *et al.*, 2006). Similar to the prior study, variant allele of *XRCC3* was related to poor survival in case of acute myeloid leukemia (Liu *et al.*, 2008). For the histology based analysis trends were observed in case of SQCC and SCLC subtype. Patients with heterozygous genotype (*CT*) in case of SQCC subtype had a better survival compared to subjects with wild type genotype (*CC*). Whereas in case of SCLC subtype subjects with wild genotype (*CC*) had a better survival compared to heterozygotes (*CT*). Osawa *et al.* in their hospital based population study revealed that SQCC population with heterozygous genotype were observed to have shorter survival time compared to wild type genotype (13 months versus 48 months) (Osawa *et al.*, 2012). Study by Butkiewicz *et al.* reported that mutant genotype (*TT*) for *XRCC3 Thr²⁴¹Met* was associated with increased risk of death (Butkiewicz *et al.*, 2011). The risk was further

observed to be higher in adenocarcinomas for mutant genotype (*TT*).

XRCC4 forms an important role in double strands breaks repair and DNA-PK, DNA-ligase and *XRCC4* are important players of NHEJ pathway. *XRCC4* is observed to act as substrate of DNA-PK since it is recruited to double strand break along with DNA-PK. It was observed that after radiation or VDJ recombination DNA-PK phosphorylations sites are not required for survival (Yu *et al.*, 2003). Our data suggested that both the DSBR genes *XRCC6 61C>G* and *XRCC7 6721G>T* were observed not to modify the prognosis for lung cancer patients as a whole in the current study. However, when stratified on the basis of histology mutant genotype (*GG*) for *XRCC6 61C>G* had a poorer OS (MST=1.4, *p*=0.0017) and a high death rate (H.R=23.9, *p*=0.0017) suggesting a poor prognosis. There is a lack of prior studies exploring the overall survival with polymorphic *XRCC6 61C>G* and *XRCC7 6721G>T* genes.

Survival of lung cancer subjects carrying polymorphic variants of BER and DSBR genes and treated with chemotherapy regimens

An important highlight of the current study is the direct implication of different chemotherapy regimens on overall survival of lung cancer subjects undergoing platinum based chemotherapy. There are very few reports on independent evaluation of chemotherapeutic drugs in relation to major polymorphic variants of the current study. There are reports of *XRCC1*, *XRCC3* in relation to certain chemotherapeutic drugs as discussed ahead. In an interesting finding, one of the previous study on lung cancer evaluated the effect of the three third generation chemotherapy regimens (cisplatin and gemcitabine, cisplatin and docetaxel, carboplatin and docetaxel) on survival. Combination of cisplatin and paclitaxel was considered as a reference for the group. The results for the study indicated no improved survival for the three regimens of third generation where as reference continued to be used in patients with good performance status and improved one year survival to 33% (Schiller *et al.*, 2002).

In the present work for subjects with polymorphic *OGG1Ser³²⁶Cys*, it was observed that patients treated with pemetrexed in combination with cisplatin/carboplatin showed a higher death rate when compared to wild type. Similarly, patients treated with irinotecan showed a better survival when compared to wild type genotype (*GC*: HR=0.051, *p*=0.0025; *GC+CC*: HR=0.056, *p*=0.0021). Further in case of *MUTYH Gln³²⁴His* it was observed that survival time was smaller for mutant type (*CC*) when compared to wild type (*GG*) (3.6 vs 10.1, *p*=0.0096) subjects where patients received docetaxel along with platinum drug. Consecutively, the univariate death rate was higher (H.R=5.34) though multivariate analysis could not solve the equation may be due to less number of sample size.

Further, our findings for the polymorphic variants of *XRCC1* suggested that patients treated with cisplatin+docetaxel/carboplatin+docetaxel and harboring the mutant genotype (AA) for *XRCC1 399* were found to have an increased OS (H.R=0.26, p=0.05) as compared to subjects with wild type genotype. Lung cancer patients who were carriers of at least single allelic variants (GA/GA+AA) and being treated with cisplatin+irinotecan / carboplatin+irinotecan were also found to have a good prognosis and an improved survival (H.R=0.22, **p=0.014**). Gastric cancer subjects receiving oxiplatin based chemotherapy in Chinese subjects exhibited a worse survival (Huang *et al.*, 2009). Whereas administration of oxiplatin as adjuvant drugs in gastric cancer subjects with polymorphic variant *XRCC1 399* revealed a better survival (Liu *et al.*, 2007). Study subjects with polymorphic *XRCC1 399* for advanced colorectal cancer suggested that polymorphic variant may be associated with resistance towards oxiplatin/5-FU (Stoehlmacher *et al.*, 2001).

On the contrary, in case of *XRCC1 Arg¹⁹⁴Trp* subjects who were heterozygous (CT) and were treated with cisplatin+docetaxel/ carboplatin+docetaxel were found to have a shorter survival rate (H.R=2.35, **p=0.012**) whereas, heterozygotes when treated with cisplatin+irinotecan/ carboplatin+irinotecan had an overall increased survival (H.R=0.34, p=0.06).

For the polymorphic variant of *XRCC3* our study reported a shorter survival for wild type genotype (CC) when treated with Cisplatin/Carboplatin+Docetaxel compared to heterozygotes (6.7 vs 9.7). On the other hand heterozygotes showed a better survival compared to wild type genotype in case of subjects treated with Cisplatin/Carboplatin+Irinotecan (10.4 vs 7.1). *XRCC3 Thr²⁴¹Met* has been studied extensively in relation to the response of gemcitabine administration. Study by Penas *et al.* revealed in their study that *XRCC3 Thr²⁴¹Met* is associated with favorable survival in patients treated with cisplatin +gemcitabine (De las Penas *et al.*, 2006). On similar lines another study showed that carriers of combined variant genotype depicted a better survival compared to wild type genotype (19.0 vs 12.5) when treated with gemcitabin+cisplatin/carboplatin whereas those treated with non-gemciatbin + cisplatin/carboplatin for the same genotype revealed a better survival for wild type win comparison to combined variant (16.0 vs 9.3) (De las Penas *et al.*, 2006). The other two DSB genes *XRCC6* and *XRCC7* were also evaluated for their independent determinance.

BER and DSB genes and their role in overall survival and objective response of lung cancer patients

The present study, also evaluated the potential of genetic variants of BER and DSB genes in predicting the objective response of lung cancer patients undergoing doublet

platinum based chemotherapy. Antineoplastic effect of platinum drugs come into activity through direct binding to DNA which results in cross-linking and ultimately apoptosis. NSCLC tumor cells exert resistance to the platinum drugs may be through detoxification, increased DNA repair capacity and intracellular accumulation. Accumulation of drugs in the form of DNA-adducts leads to the cytotoxicity of the drug. Clinical outcome or response towards chemotherapy is related to platinum DNA-adducts in blood. Higher DNA-adducts corresponds to better response towards cancer. Polymorphic DNA repair genes would result in decreased removal of platinum DNA-adducts in circulation. Therefore in this sequence efforts have been made to find out the plausible role of BER and DSBR genes in altering the response of patients towards chemotherapy in North Indian population. In the current study *OGG1* was observed to positively influence the response different from a Chinese study where no association was observed between SNP and response rate. Whereas, similar to our study resistance rate was higher for heterozygous genotype compared to wild subjects in the study (Peng *et al.*, 2014). The efficacy of the chemotherapy treatment is observed through response rate towards chemotherapy regimens administered.

For the polymorphic variants of *XRCC1* our study reported an unfavorable association of *XRCC1 399* with response (*GA*, OR=0.74; *GA+AA*, OR=0.92). Studies in Chinese subject also reported the similar results (Zhao and Chen, 2015; Hong *et al.*, 2009). Meta-analytical study also concluded carriers of at least single variant allele (*GA/GA+AA*) depicted an unfavourable response towards chemotherapy (Yuan *et al.*, 2015). However, study by Z.Y Jin *et al.* (Jin *et al.*, 2014) reported a good response of *XRCC1 399* mutant carriers (*AA*) and no association for *XRCC1 194*. Further, studies have reported *XRCC1 194* is highly associated with good response (Cao *et al.*, 2014; Zhao and Chen, 2015; Yuan *et al.*, 2015) whereas our study showed no association with response towards chemotherapy for the SNP. Heterozygotes (*GA*) and combined variants (*GG +GA*) for *XRCC1 280* were observed to have no effect on response similar to two Chinese studies (Jin *et al.*, 2014; Yuan *et al.*, 2015). However mutant genotype (*AA*) was observed to be categorized as poor responders (OR=0.32) contrary to the work of Zhao *et.al* (Zhao and Chen, 2015).

The evaluation of DSBR genes with objective response revealed that *XRCC3* did not show any enhancement in response, similar to our study genotypes of *XRCC3* in Chinese study were not associated with tumor response (Chen, Sun, Ren, Curran, Zhang, Zhou, Zhang and Zhou, 2012). However, a metanalytical study, depicted the involvement of the variant allele with good response towards chemotherapy. The study for stratified subgroups analysis exhibited a good response for Caucasian and mixed population but

not for Asian subjects (Shen *et al.*, 2013). A positive response was observed for the next DSB repair gene *i.e.* *XRCC4* for subjects carrying variant genotype compared to subjects with wild type genotype. There is no study till date, with such clinical implications of the genetic variants.

6.1.5 Gene-gene interactions in lung cancer predisposition and survival

Gene-gene interaction in lung cancer predisposition

DNA damage is a serious threat to the organisms, which question the complete transfer of genetic material to the progeny as well as the survival of organisms. To counteract these threats organisms have developed DNA damage responses. DNA damage response is a complex pathway which sense damages, transduce the signals to effectors and provoke cellular response towards damages. DNA repair is carried out by a plethora of enzymes that chemically modulate DNA when repairing dna damage. DNA repair and carcinogenesis has been immensely explored. Cancer is multifactorial disease that involve complex interactions between multiple genes. Individual allele is considered to have low penetrance and cannot be considered responsible for multiple gene interactions, which justify the inconsistency in various cancer studies. Therefore, the present study has made an attempt to unfold the hidden interactive contributions conferred by the panel of genetic variants of DNA repair pathways in lung cancer predisposition. To overcome the shortcomings of single gene based approach, multi-gene interactions are analyzed in the present study. This is probably the first attempt to evaluate the collaborative effect of the polymorphic variants in DNA repair genes on lung cancer risk and findings are new and have not been found in large scale investigations. This can add to our knowledge of lung cancer predisposition.

The double combinations analyzed in the study revealed that lung cancer patients having variant or combined variant for the combinations of *OGG1 Ser³²⁶Cys* and *XRCC4*, *MutYh Gln³²⁴His* and *XRCC4*, *OGG1 Ser³²⁶Cys* and *XRCC7*, *MutYh Gln³²⁴His* and *XRCC6*, *MutYh Gln³²⁴His* and *XRCC7*, *XRCC1 Arg³⁹⁹Gln* and *OGG1 Ser³²⁶Cys*, *XRCC1 Arg¹⁹⁴Trp* and *OGG1 Ser³²⁶Cys*, *XRCC1 Pro²⁰⁶Pro* and *OGG1 Ser³²⁶Cys*, *XRCC1 Pro²⁰⁶Pro* and *MutYh Gln³²⁴His* and for every double combination of *XRCC1 Gln⁶³²Gln* can be classified as high risk subgroup. Certain combinations such as *XRCC1 Arg²⁸⁰His* and *OGG1 Ser³²⁶Cys*, *XRCC1 Arg²⁸⁰His* and *XRCC6 61C>G*, *XRCC1 Arg²⁸⁰His* and *XRCC7* were found to be categorized as low risk subgroups. There have been very few studies which have assessed the cumulative power of genes together. An Indian study in bladder cancer

exhibited an elevated risk for both variant and combined variant genotype for the combination of *OGG1 Ser³²⁶Cys* and *XRCC7 6721 G>T* similar to the present study (Gangwar *et al.*, 2009). Whereas, findings from another study in bladder cancer for the double combination of *OGG1 Ser³²⁶Cys* and *XRCC1 Arg³⁹⁹Gln* having *CG* of *OGG1 Ser³²⁶Cys* and *AA* of *XRCC1 Arg³⁹⁹Gln* classified the group as low risk subgroup in contrary to our study in lung cancer which classified the group as a high risk subgroup (Srivastava *et al.*, 2009). The heterozygous subjects for the combination of *OGG1 Ser³²⁶Cys* and *XRCC1 Arg³⁹⁹Gln* in glaucoma similar to the present work exhibited a high risk (Szaflik *et al.*, 2013). Another study has investigated the combination of *OGG1 Ser³²⁶Cys* and *XRCC1 Arg³⁹⁹Gln* for subjects carrying single allele of variant genotype in lung cancer and similar to our work the heterozygous subjects were at high risk (Letkova *et al.*, 2013). Study in colorectal cancer also reported similar results (Przybyłowska *et al.*, 2013). In an study in Chinese population in lung cancer the dichotomous group for the combination of *XRCC1 399*, *OGG1* and *MUTYH* revealed a high risk (Qian *et al.*, 2011). Study in colorectal cancer also reported the similar results (Przybyłowska *et al.*, 2013). The interactive potential of sequence variants *XRCC1 Arg³⁹⁹Gln* and *MUTYH* for the study revealed a high risk, the above double combination for subjects carrying single allelic variants in glaucoma exhibited a high risk (Szaflik *et al.*, 2013). The current study also evaluated the double combinations for inter-pathway genes, in this sequence the interaction between *XRCC1 Arg³⁹⁹Gln* and *XRCC7* revealed no risk for the combination. There are no reports on lung cancer for such interactions however, a study in bladder cancer evaluated the impact of the double combination of *XRCC1 Arg³⁹⁹Gln* and *MUTYH* and a high risk was observed for the variant genotype (Zhi *et al.*, 2012).

The genetic combination of namely *GG* genotype of *OGG1 Ser³²⁶Cys* and *CT* genotype of *XRCC1 Arg¹⁹⁴Trp* suggested a high risk towards bladder cancer (Srivastava *et al.*, 2009). Similarly, investigations in colorectal cancer for the double combination having *CG* of *OGG1 Ser³²⁶Cys* and *CC* of *XRCC1 Arg¹⁹⁴Trp* possessed a high risk (Przybyłowska *et al.*, 2013).

Several genetic variants in the study have been evaluated together in different ethnicities modulating the risk of cancers. Indian study on breast cancer evaluated *XRCC4*, *XRCC6*, *XRCC7*. A lack of association was observed for *XRCC4* while a protective effect was observed in *XRCC6* variant and a high risk was observed in polymorphic variants of *XRCC7* (Patil *et al.*, 2016). Another study in head and neck cancer depicted an interesting finding where subjects with at-least one variant allele *i.e.* heterozygotes or double allelic variants for *XRCC1 Arg¹⁹⁴Trp*, *XRCC6* and every genotypic combination for *XRCC7*, *XRCC1 Arg³⁹⁹Gln* was found to be positively correlated with head and neck

cancer (Datkhile *et al.*, 2016).

The above studies have drawn conclusions on the basis of single gene. There are very rare reports on collaborative effect of DNA repair genes on modulation of lung cancer. There are studies which have looked into interactive potential of several genetic variants in various cancer such as bladder (Huang *et al.*, 2007; Horikawa *et al.*, 2008), gall bladder (Rai *et al.*, 2015) head and neck cancer (Das *et al.*, 2015), breast cancer (Ritchie *et al.*, 2001), gastric adenocarcinoma (Liu *et al.*, 2010). These studies have employed a high order statistics which has high power in forecasting the multiple-faceted role of these SNPs along with their interaction with the environmental parameters such as smoking. MDR and CART tools were used and some interesting conclusions about the contributions of all the germline variations under study became evident. These non-parametric methods have gained much more significance recently as they give us a clear picture about the various factors which might play an inevitable role in impeding cancer predisposition.

MDR and CART analysis in genetic variants of BER and DSBR genes

Interesting finding of the present study comprised the role of *XRCC1 Pro²⁰⁶Pro*, *XRCC1 Gln⁶³²Gln*, *OGG1 Ser³²⁶Cys*, along with *XRCC1 Arg²⁸⁰His* in lung cancer predisposition. These sequence variants contributed towards increasing the risk for lung cancer. This is the first study to evaluate the BER and DSBR genes and their cumulative effect on lung cancer risk. Interactions model within repair pathway showed the evident partnership of *XRCC1 Pro²⁰⁶Pro*, *XRCC1 Gln⁶³²Gln* and *OGG1 Ser³²⁶Cys*. A Chinese study in lung cancer has observed that *XRCC1 Arg¹⁹⁴Trp*, *XPC PAT*, *FAS G-1377A*, *FASL T-844C* was taken as a parameter to find out the the best interaction model (Liu *et al.*, 2013). Another study in bladder cancer used multidimensional MDR method in bladder cancer to find out the interactions between smoking status, *OGG1*, *APEX1*, *ADPRT*, *MUTYH* and *XRCC1* (Huang *et al.*, 2007). In an Indian study in head and neck cancer MDR tool is used to find the best interaction model in different SNPs of *XPB*, *APE1*, *MUTYH* (Das *et al.*, 2015).

Considering the CART results in case of repair genes *XRCC1 632* came out to be the most significant factor in determining the lung cancer risk. *XRCC1 632* interacted with other DNA repair variants of *OGG1*, *MUTYH*, *XRCC4*, *XRCC6*, *XRCC7*. When gene environment interactions were taken into consideration along with polymorphic variants smoking formed the initial split indicating smokers as a significant factor. Similar to overall CART, it was identified in smokers that *XRCC1 632* was the most important factor. The role of DNA repair pathways was illustrated by genotypic combinations such as *XRCC1 632(M)*, *MUTYH 324(M)*, *XRCC1 194(W)*, *XRCC4(M)*, *XRCC1 280(W)* where a high risk for lung cancer was indicated.

Ours is a primary study making use of polygenic approach to obtain evidence about repair genes in determining lung cancer susceptibility in North Indian population. MDR is used to find the non-linear interactions for the study which also involve synonymous polymorphisms of *XRCC1* which are evaluated for first time in lung cancer. The presence of single SNPs in more than one model indicate the strong contribution of the SNP in lung cancer predisposition. This hints to further investigate the SNPs with large sample size in different ethnicities.

Gene-gene interaction in lung cancer survival

An attempt was made to look for the double combinations of genetic variants of the study modifying the survival of lung cancer subjects. It became evident from this examination that major double combinations did not influence the survival of subjects whereas subjects for the combination of *XRCC1 Arg²⁸⁰His* and *OGG1 Ser³²⁶Cys* exhibited a small survival time (16.1 vs 6.8 months) with high death rate (H.R=2.21, $p=0.05$). A Chinese study in lung cancer with increasing variant alleles for the combination of *OGG1 Ser³²⁶Cys* and *XRCC1 Arg³⁹⁹Gln* had a high risk to die early in short period (H.R=1.9, $p\leq 0.001$) (Peng *et al.*, 2014). Another study in lung cancer where *XRCC1 Arg³⁹⁹Gln* with *XPD* similar to the above report stated a high hazard rate with increasing variant alleles (Gurubhagavatula *et al.*, 2004). Introspection in bladder cancer in an Italy based study analyzed the combined effect of certain polymorphic variants of *XRCC1*. The report revealed that with increasing variant allele the survival rate increased and a 26% decrease in death rate was observed for every added variant allele (Sacerdote *et al.*, 2013). Combination of genetic variants of *XRCC1*, *XRCC3*, *CCND* in breast cancer elevated the hazard rate with increasing variant allele (Bewick *et al.*, 2006).

Another important highlight of the present study is use of high order interaction tools to find out the subgroups showing the best prognosis and higher OS due to the presence of the different genotypic combinations which cannot be inferred by traditional statistical tools. This was probably the first attempt to analyze the interaction of DNA repair genes in predicting the survival of lung cancer patients depicted by a survival tree where *OGG1 Ser³²⁶Cys* came out to be the most significant factor in modulating the survival of lung cancer subjects. Node 7 *OGG1 326(M)/XRCC3 (M)/XRCC1 194(W)/MUTYH(M)/XRCC1 399(M)/XRCC1 206(M)/XRCC6(W)* revealed the lowest death rate (H.R=0.073, $p=0.02$). Another study done in Chinese population on lung cancer explored the interactive potential of *XRCC1 Arg³⁹⁹Gln*, *OGG1 Ser³²⁶Cys*, *MUTYH Gln³²⁴Cys*, *XPA (rs1800975) G>A*, *ERCC2 Asp³¹²Asn* and *XRCC3 Thr²⁴¹Met* polymorphism using recursive partitioning approach to obtain a decision tree to predict the survival (Su *et al.*, 2015). Studies in never smokers in NSCLC subjects have used

recursive partitioning by S-tree to find the high risk sequence in lung cancer survival (Pu *et al.*, 2012).

In conclusion, this study revealed the association of polymorphic variants of DNA repair genes in survival of North Indian lung cancer subjects. However, limitation to the above study is the small sample size, and lack of functional studies. The analysis for the stratified subgroups uncovers various the role of factors playing important part in lung cancer progression. In order to validate the above findings investigation in large sample size is warranted.

Chapter 7

Conclusions

Findings from the present study

- Lung cancer risk increased two times in subjects with CG genotype for *OGG1 Ser³²⁶Cys*, whereas a highly protective effect was observed in subjects with GG genotype
- The subjects with heterozygous genotype *MUTYH Gln³²⁴His* were found to have a higher risk of developing lung cancer. The single allelic variant for the genotypic combination of *OGG1 Ser³²⁶Cys* and *MUTYH Gln³²⁴His* showed a seven-fold increased risk of developing lung cancer
- *XRCC1 Arg³⁹⁹Gln* did not confer any effect on lung cancer susceptibility
- *XRCC1 Gln⁶³²Gln* is the best one factor model determining lung cancer risk as determined by MDR. Second best model was three factor model comprising of *XRCC1 Gln⁶³²Gln*, *XRCC1 Pro²⁰⁶Pro*, *XRCC1 Gln²⁸⁰His*
- The polymorphic variants of DSBR genes *XRCC4*, *XRCC6*, *XRCC7* did not modulate lung cancer risk.
- Patients with genotypic combination of terminal node 1 of CART tree (*XRCC1 Gln⁶³²Gln*(H+M), *XRCC1 Pro²⁰⁶Pro*(W)+ *MUTYH Gln³²⁴His*) showed an OR of 29 fold.
- *XRCC1 Gln⁶³²Gln* did show a high death rate for SCLC subjects
- Lung cancer patients administered Cisplatin/Carboplatin+Irinotecan and Cisplatin/Carboplatin+Pemetrexed for polymorphic *OGG1 Ser³²⁶Cys* did show a poor survival.
- *XRCC1 Arg³⁹⁹Gln* and *XRCC1 Arg¹⁹⁴Trp* did show a better survival for subjects treated with Cisplatin/Carboplatin+Irinotecan.
- It was observed unfavourable genotype of the terminal node 7 of S-tree (*OGG1 326(M)*/*XRCC3 (M)*/*XRCC1 194(W)*/*MUTYH(M)*/*XRCC1 399(M)*/*XRCC1 206(M)*/*XRCC6(W)*) exhibited a highly decreased rate of death. Survival tree

analysis predicted *OGG1 326(W)* as the main contributor in survival of lung cancer patients.

In conclusion this is the first Indian study to highlight such large number of DNA repair genes altogether with such a large sample size and also their correlation with clinicopathological factors. The stratified analysis was also performed for smoking and histology to find out high risk posing subgroups. Subjects with single allelic variant for initial BER genes *OGG1* and *MUTYH* and *XRCC1 Gln⁶³²Gln* were observed to be positively correlated with lung cancer predisposition. Our study is an first attempt in DNA repair genes for lung cancer in Indian population to make use of multifactorial approach (CART and MDR) to explore the complex interaction between the DNA repair genes and its relationship towards lung cancer risk. This is the first study in the North Indian region to analyze the SNPs together for overall survival and clinical outcomes. We for the first time have analyzed the stratified clinical outcomes with genotype for the genetic variants in the study. The present findings with larger sample size for overall and stratified groups would help in development of bio-markers both for diagnosis of lung cancer and prognosis of lung cancer.

7.1 Future Prospects of the study

DNA repair enzymes are correctives for DNA damage induced by carcinogens and anti-cancer drug, it is very likely that SNPs in DNA repair genes may play an important part in cancer susceptibility. These polymorphisms in the DNA repair pathways are worth paying more attention for it may serve as a platform for detecting and predicting individual susceptibility for lung cancer. Secondly it might also help to reveal the prognosis or the disease. To fulfill the individualized therapy system, the construction of comprehensive and reliable SNP and haplotype analyses of different ethnic population like Indian will be very useful and shall facilitate predictive genotype-drug efficacy in each individual patient. Data suggests that several SNPs in the DNA repair system contribute to inter individual variability and, therefore, could be valuable predictive factors for the development of personalized drugs

References

- Abdel-Rahman, S. Z. and El-Zein, R. A. (2000), ‘The 399gln polymorphism in the dna repair gene xrccl modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine nnk’, *Cancer letters* **159**(1), 63–71.
- Al-Abd, A. M., Alamoudi, A. J., Abdel-Naim, A. B., Neamatallah, T. A. and Ashour, O. M. (2017), ‘Anti-angiogenic agents for the treatment of solid tumors: Potential pathways, therapy and current strategies—a review’, *Journal of advanced research* **8**(6), 591–605.
- Almeida, K. H. and Sobol, R. W. (2007), ‘A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification’, *DNA repair* **6**(6), 695–711.
- American Cancer Society (2017).
URL: <https://www.cancer.org/cancer/non-small-cell-lung-cancer/detection-diagnosis-staging/staging.html>
- Ang, M.-K., Patel, M. R., Yin, X.-Y., Sundaram, S., Fritchie, K., Zhao, N., Liu, Y., Freemerman, A. J., Wilkerson, M. D., Walter, V. *et al.* (2011), ‘High xrccl protein expression is associated with poorer survival in patients with head and neck squamous cell carcinoma’, *Clinical Cancer Research* **17**(20), 6542–6552.
- Ankathil, R. (2010), ‘Tobacco, genetic susceptibility and lung cancer’, *Tobacco Use Insights* **3**, TUI-S2819.
- Ansari, J., Palmer, D., Rea, D. and Hussain, S. (2009), ‘Role of tyrosine kinase inhibitors in lung cancer’, *Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents)* **9**(5), 569–575.
- Asthana, S., Labani, P. S., Labani, S. *et al.* (2015), ‘A review on cancer incidence in india from 25 population-based cancer registries’, *Journal of Dr. NTR University of Health Sciences* **4**(3), 150.
- Au, W. W., Navasumrit, P. and Ruchirawat, M. (2004), ‘Use of biomarkers to characterize functions of polymorphic dna repair genotypes’, *International journal of hygiene and environmental health* **207**(4), 301–313.
- Au, W. W., Salama, S. A. and Sierra-Torres, C. H. (2003), ‘Functional characterization of polymorphisms in dna repair genes using cytogenetic challenge assays.’, *Environmental health perspectives* **111**(15), 1843.
- Bahl, A. and Falk, S. (2001), ‘Meta-analysis of single agents in the chemotherapy of nslc: what do we want to know?’, *British journal of cancer* **84**(9), 1143.

- Bau, D.-T., Tsai, C.-W. and Wu, C.-N. (2011), 'Role of the xrc5/xrc6 dimer in carcinogenesis and pharmacogenomics', *Pharmacogenomics* **12**(4), 515–534.
- Beard, W. A., Prasad, R. and Wilson, S. H. (2006), 'Activities and mechanism of dna polymerase β ', *Methods in enzymology* **408**, 91–107.
- Beggs, A. D., Domingo, E., McGregor, M., Presz, M., Johnstone, E., Midgley, R., Kerr, D., Oukrif, D., Novelli, M., Abulafi, M. *et al.* (2012), 'Loss of expression of the double strand break repair protein atm is associated with worse prognosis in colorectal cancer and loss of ku70 expression is associated with cin', *Oncotarget* **3**(11), 1348.
- Behera, D. (2007), 'New approach to the treatment of lung cancer: the molecular targeted therapy', *INDIAN JOURNAL OF CHEST DISEASES AND ALLIED SCIENCES* **49**(3), 149.
- Behera, D. and Balamugesh, T. (2004), 'Lung cancer in india', *Indian Journal of Chest Diseases and Allied Sciences* **46**, 269–282.
- Benjamini, Y. and Hochberg, Y. (1995), 'Controlling the false discovery rate: a practical and powerful approach to multiple testing', *Journal of the royal statistical society. Series B (Methodological)* pp. 289–300.
- Berhane, N., Sobti, R. C. and Mahdi, S. A. (2012), 'Dna repair genes polymorphism (xpg and xrc1) and association of prostate cancer in a north indian population', *Molecular biology reports* **39**(3), 2471–2479.
- Berman, D. W. and Crump, K. S. (2008), 'Update of potency factors for asbestos-related lung cancer and mesothelioma', *Critical reviews in toxicology* **38**(sup1), 1–47.
- Bewick, M. A., Conlon, M. S. and Lafrenie, R. M. (2006), 'Polymorphisms in xrc1, xrc3, and ccnd1 and survival after treatment for metastatic breast cancer', *Journal of clinical oncology* **24**(36), 5645–5651.
- Birben, E., Sahiner, U. M., Sackesen, C., Erzurum, S. and Kalayci, O. (2012), 'Oxidative stress and antioxidant defense', *World Allergy Organization Journal* **5**(1), 9.
- Brenneman, M. A., Weiss, A. E., Nickoloff, J. A. and Chen, D. J. (2000), 'Xrc3 is required for efficient repair of chromosome breaks by homologous recombination', *Mutation Research/DNA Repair* **459**(2), 89–97.
- Brooks, D. R., Austin, J. H., Heelan, R. T., Ginsberg, M. S., Shin, V., Olson, S. H., Muscat, J. E. and Stellman, S. D. (2005), 'Influence of type of cigarette on peripheral versus central lung cancer', *Cancer Epidemiology and Prevention Biomarkers* **14**(3), 576–581.
- Brown, J., Eraut, D., Trask, C. and Davison, A. (1996), 'Age and the treatment of lung cancer.', *Thorax* **51**(6), 564–568.
- Budden, T. and Bowden, N. A. (2013), 'The role of altered nucleotide excision repair and uvb-induced dna damage in melanomagenesis', *International Journal of Molecular Sciences* **14**(1), 1132–1151.

- Butkiewicz, D., Rusin, M., Sikora, B., Lach, A. and Chory, M. (2011), ‘An association between dna repair gene polymorphisms and survival in patients with resected non-small cell lung cancer’, *Molecular biology reports* **38**(8), 5231–5241.
- Cabanero, M., Sangha, R., Sheffield, B., Sukhai, M., Pakkal, M., Kamel-Reid, S., Karsan, A., Ionescu, D., Juergens, R., Butts, C. *et al.* (2017), ‘Management of egfr-mutated non-small-cell lung cancer: practical implications from a clinical and pathology perspective’, *Current Oncology* **24**(2), 111.
- Cadet, J. and Wagner, J. R. (2013), ‘Dna base damage by reactive oxygen species, oxidizing agents, and uv radiation’, *Cold Spring Harbor perspectives in biology* **5**(2), a012559.
- Caldecott, K. W. (2003), ‘Xrcc1 and dna strand break repair’, *DNA repair* **2**(9), 955–969.
- Caldecott, K. W. (2008), ‘Single-strand break repair and genetic disease’, *Nature Reviews Genetics* **9**(8), 619.
- Caldecott, K. W., Aoufouchi, S., Johnson, P. and Shall, S. (1996), ‘Xrcc1 polypeptide interacts with dna polymerase β and possibly poly (adp-ribose) polymerase, and dna ligase iii is a novel molecular nick-sensor in vitro’, *Nucleic acids research* **24**(22), 4387–4394.
- Cancer Facts & Figures* (2016).
- URL:** <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2016.html>
- Cao, Z., Song, J., Wang, J., Guo, X., Yu, S. and Dong, W. (2014), ‘Association between polymorphisms in xrcc1 gene and treatment outcomes of patients with advanced gastric cancer: a systematic review and meta-analysis’, *PloS one* **9**(1), e85357.
- Cătană, A., Pop, M., Hincă, B. D., Pop, I. V., Petrișor, F. M., Porojan, M. D. and Popp, R. A. (2015), ‘The xrcc1 arg194trp polymorphism is significantly associated with lung adenocarcinoma: a case-control study in an eastern european caucasian group’, *OncoTargets and therapy* **8**, 3533.
- Caulfield, J. L., Wishnok, J. S. and Tannenbaum, S. R. (1998), ‘Nitric oxide-induced deamination of cytosine and guanine in deoxynucleosides and oligonucleotides’, *Journal of Biological Chemistry* **273**(21), 12689–12695.
- Cecchin, E., Agostini, M., Pucciarelli, S., De Paoli, A., Canzonieri, V., Sigon, R., De Mattia, E., Friso, M., Biason, P., Visentin, M. *et al.* (2011), ‘Tumor response is predicted by patient genetic profile in rectal cancer patients treated with neo-adjuvant chemoradiotherapy’, *The pharmacogenomics journal* **11**(3), 214.
- Chacko, P., Rajan, B., Joseph, T., Mathew, B. S. and Pillai, M. R. (2005), ‘Polymorphisms in dna repair gene xrcc1 and increased genetic susceptibility to breast cancer’, *Breast cancer research and treatment* **89**(1), 15–21.
- Chan, D. W. and Lees-Miller, S. P. (1996), ‘The dna-dependent protein kinase is inacti-

- vated by autophosphorylation of the catalytic subunit', *Journal of Biological Chemistry* **271**(15), 8936–8941.
- Chang (2008), 'Nucleotide excision repair genes and risk of lung cancer among san francisco bay area latinos and african americans', *International journal of cancer* **123**(9), 2095–2104.
- Chang, A. (2011), 'Chemotherapy, chemoresistance and the changing treatment landscape for nsclc', *Lung cancer* **71**(1), 3–10.
- Chang, C.-H. and Chang, C.-L. (2009), 'Significant association of an xrcc4 single nucleotide polymorphism with bladder cancer susceptibility in taiwan', *Anticancer Research* **29**(5), 1777–1782.
- Chang, C. H. and Hsiao, C. F. (2009), 'Interactive effect of cigarette smoking with human 8-oxoguanine dna n-glycosylase 1 (hogg1) polymorphisms on the risk of lung cancer: a case-control study in taiwan', *American journal of epidemiology* **170**(6), 695–702.
- Chang, C. M., Edwards, S. H., Arab, A., Del Valle-Pinero, A. Y., Yang, L. and Hatsukami, D. K. (2017), 'Biomarkers of tobacco exposure: summary of an fda-sponsored public workshop', *Cancer Epidemiology and Prevention Biomarkers* **26**(3), 291–302.
- Chang, J. S., Wrensch, M. R., Hansen, H. M., Sison, J. D., Aldrich, M. C., Quesenberry Jr, C. P., Seldin, M. F., Kelsey, K. T. and Wiencke, J. K. (2008), 'Base excision repair genes and risk of lung cancer among san francisco bay area latinos and african-americans', *Carcinogenesis* **30**(1), 78–87.
- Chang, W.-S., Ke, H.-L., Tsai, C.-W., Lien, C.-S., Liao, W.-L., Lin, H.-H., Lee, M.-H., Wu, H.-C., Chang, C.-H., Chen, C.-C. *et al.* (2012), 'The role of xrcc6 t-991c functional polymorphism in renal cell carcinoma', *Anticancer research* **32**(9), 3855–3860.
- Chen, J., Zhao, Q.-w. and Shi, G.-m. (2012), 'Xrcc1 arg399gln and clinical outcome of platinum-based treatment for advanced non-small cell lung cancer: a meta-analysis in 17 studies', *Journal of Zhejiang University SCIENCE B* **13**(11), 875–883.
- Chen, L., Zhuo, D., Chen, J. and Yuan, H. (2015), 'Xrcc1 polymorphisms and lung cancer risk in caucasian populations: a meta-analysis', *International journal of clinical and experimental medicine* **8**(9), 14969.
- Chen, M., Kamat, A. M., Huang, M., Grossman, H. B., Dinney, C. P., Lerner, S. P., Wu, X. and Gu, J. (2007), 'High-order interactions among genetic polymorphisms in nucleotide excision repair pathway genes and smoking in modulating bladder cancer risk', *Carcinogenesis* **28**(10), 2160–2165.
- Chen, X., Sun, H., Ren, S., Curran, V. K., Zhang, L., Zhou, S., Zhang, J. and Zhou, C. (2012), 'Association of xrcc3 and xpd751 snp with efficacy of platinum-based chemotherapy in advanced nsclc patients', *Clinical and Translational Oncology* **14**(3), 207–213.

- Cheng, C.-X., Xue, M., Li, K. and Li, W.-S. (2012), 'Predictive value of *xrcc1* and *xrcc3* gene polymorphisms for risk of ovarian cancer death after chemotherapy', *Asian Pacific journal of cancer prevention* **13**(6), 2541–2545.
- Chiu, C.-F., Tsai, M.-H., Tseng, H.-C., Wang, C.-L., Wang, C.-H., Wu, C.-N., Lin, C.-C. and Bau, D.-T. (2008), 'A novel single nucleotide polymorphism in *xrcc4* gene is associated with oral cancer susceptibility in taiwanese patients', *Oral oncology* **44**(9), 898–902.
- Cho, E.-Y., Hildesheim, A., Chen, C.-J., Hsu, M.-M., Chen, I.-H., Mittl, B. F., Levine, P. H., Liu, M.-Y., Chen, J.-Y., Brinton, L. A. *et al.* (2003), 'Nasopharyngeal carcinoma and genetic polymorphisms of dna repair enzymes *xrcc1* and *hogg1*', *Cancer Epidemiology and Prevention Biomarkers* **12**(10), 1100–1104.
- Chu, P. G. and Weiss, L. M. (2002), 'Expression of cytokeratin 5/6 in epithelial neoplasms: an immunohistochemical study of 509 cases', *Modern pathology* **15**(1), 6.
- Ciccia, A. and Elledge, S. J. (2010), 'The dna damage response: making it safe to play with knives', *Molecular cell* **40**(2), 179–204.
- Cimino, G. D., Pan, C.-x. and Henderson, P. T. (2013), 'Personalized medicine for targeted and platinum-based chemotherapy of lung and bladder cancer', *Bioanalysis* **5**(3), 369–391.
- Clancy, S. (2008), 'Dna damage & repair: mechanisms for maintaining dna integrity', *Nature Education* **1**(1), 103.
- Cohen, S. M. and Arnold, L. L. (2010), 'Chemical carcinogenesis', *Toxicological Sciences* **120**(suppl_1), S76–S92.
- Colby, T. V. (1995), 'Tumors of the lower respiratory tract', *Atlas of tumor pathology* pp. 309–317.
- Control, U. D. and Prevention (2010), *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General*, National Center for Chronic Disease Prevention and Health Promotion (US); Office on Smoking and Health (US).
- Cooper, W. A., Lam, D. C., OToole, S. A. and Minna, J. D. (2013), 'Molecular biology of lung cancer', *Journal of thoracic disease* **5**(Suppl 5), S479.
- Coté, M. L., Liu, M., Bonassi, S., Neri, M., Schwartz, A. G., Christiani, D. C., Spitz, M. R., Muscat, J. E., Rennert, G., Aben, K. K. *et al.* (2012), 'Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the international lung cancer consortium', *European Journal of Cancer* **48**(13), 1957–1968.
- Couraoud, S., Zalcman, G., Milleron, B., Morin, F. and Souquet, P.-J. (2012), 'Lung cancer in never smokers—a review', *European journal of cancer* **48**(9), 1299–1311.
- Couto, P. G., Bastos-Rodrigues, L., Carneiro, J. G., Guieiro, F., Bicalho, M. A., Leidenz,

- F. B., Bicalho, A. J., Friedman, E. and De Marco, L. (2015), 'Dna base-excision repair genes ogg1 and nth1 in brazilian lung cancer patients', *Molecular diagnosis & therapy* **19**(6), 389–395.
- Cruz, C. S. D., Tanoue, L. T. and Matthay, R. A. (2011), 'Lung cancer: epidemiology, etiology, and prevention', *Clinics in chest medicine* **32**(4), 605–644.
- Cui, Z., Yin, Z., Li, X., Wu, W., Guan, P. and Zhou, B. (2012), 'Association between polymorphisms in xrcc1 gene and clinical outcomes of patients with lung cancer: a meta-analysis', *BMC cancer* **12**(1), 71.
- Curiosity (2013), 'Radiation effects on cells and dna'.
- Dalavanga, Y. (2017), 'The principles governing the lung structure and function', *Anatomy Physiology & Biochemistry International Journal* **2**.
- Das, S., Bhowmik, A., Bhattacharjee, A., Choudhury, B., Naiding, M., Laskar, A. K., Ghosh, S. K. and Choudhury, Y. (2015), 'Xpd, ape1, and mutyh polymorphisms increase head and neck cancer risk: effect of gene-gene and gene-environment interactions', *Tumor Biology* **36**(10), 7569–7579.
- Dasika, G. K., Lin, S.-C. J., Zhao, S., Sung, P., Tomkinson, A. and Lee, E. Y. P. (1999), 'Dna damage-induced cell cycle checkpoints and dna strand break repair in development and tumorigenesis', *oncogene* **18**(55), 7883.
- Datkhile, K. D., Patil, M. N., Khamkar, T. S., Vhaval, R. D., Durgawale, P. P., Chougule, P. G., Gudur, A. and Kakade, S. V. (2017), 'Genetic polymorphisms in dna repair genes (hogg1 & ape1) and their association with oral cancer susceptibility in rural indian population: a hospital based case-control study', *International Journal of Medical Science and Public Health* **6**(1), 23–28.
- Datkhile, K., Vhaval, R., Patil, M., Khamkar, T., Durgavale, P., Gudur, A. *et al.* (2016), 'Role of genetic polymorphisms in dna repair genes (xrcc1, xrcc2, xrcc3, xrcc4, xrcc5, xrcc6, xrcc7) in head and neck cancer susceptibility in rural indian population: A hospital based case-control study from south-western maharashtra', *Int. J. Curr. Res* **8**, 25482–92.
- David-Beabes, G. L. and London, S. J. (2001), 'Genetic polymorphism of xrcc1 and lung cancer risk among african-americans and caucasians', *Lung Cancer* **34**(3), 333–339.
- Davis, A. J. and Chen, D. J. (2013), 'Dna double strand break repair via non-homologous end-joining', *Translational cancer research* **2**(3), 130.
- de Groot, P. and Munden, R. F. (2012), 'Lung cancer epidemiology, risk factors, and prevention', *Radiologic Clinics* **50**(5), 863–876.
- De las Penas, R., Sanchez-Ronco, M., Alberola, V., Taron, M., Camps, C., Garcia-Carbonero, R., Massuti, B., Queralt, C., Botia, M., Garcia-Gomez, R. *et al.* (2006), 'Polymorphisms in dna repair genes modulate survival in cisplatin/gemcitabine-treated

- non-small-cell lung cancer patients', *Annals of Oncology* **17**(4), 668–675.
- De Ruyck, K., Szaumkessel, M., De Rudder, I., Dehoorne, A., Vral, A., Claes, K., Velghe, A., Van Meerbeeck, J. and Thierens, H. (2007), 'Polymorphisms in base-excision repair and nucleotide-excision repair genes in relation to lung cancer risk', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **631**(2), 101–110.
- De Zio, D., Cianfanelli, V. and Cecconi, F. (2013), 'New insights into the link between dna damage and apoptosis', *Antioxidants & redox signaling* **19**(6), 559–571.
- Deligezer, U. and Dalay, N. (2004), 'Association of the xrccl gene polymorphisms with cancer risk in turkish breast cancer patients', *Experimental & molecular medicine* **36**(6), 572.
- Desouky, O., Ding, N. and Zhou, G. (2015), 'Targeted and non-targeted effects of ionizing radiation', *Journal of Radiation Research and Applied Sciences* **8**(2), 247–254.
- Dexheimer, T. S. (2013), Dna repair pathways and mechanisms, in 'DNA repair of cancer stem cells', Springer, pp. 19–32.
- Ding, X., Wang, K., Wu, Z., Yao, A., Li, J., Jiao, C., Qian, J., Bai, D. and Li, X. (2015), 'The ser326cys polymorphism of hogg1 is associated with intrahepatic cholangiocarcinoma susceptibility in a chinese population', *International journal of clinical and experimental medicine* **8**(9), 16294.
- Divine, K. K., Gilliland, F. D., Crowell, R. E., Stidley, C. A., Bocklage, T. J., Cook, D. L. and Belinsky, S. A. (2001), 'The xrccl 399 glutamine allele is a risk factor for adenocarcinoma of the lung', *Mutation Research/DNA Repair* **461**(4), 273–278.
- Dixon, K. and Koprass, E. (2004), Genetic alterations and dna repair in human carcinogenesis, in 'Seminars in cancer biology', Vol. 14, Elsevier, pp. 441–448.
- Downward, J. (2003), 'Targeting ras signalling pathways in cancer therapy', *Nature Reviews Cancer* **3**(1), 11.
- Du, Y., He, Y., Mei, Z., Qian, L., Shi, J. and Jie, Z. (2016), 'Association between genetic polymorphisms in xpd and xrccl genes and risks of non-small cell lung cancer in east chinese han population', *The clinical respiratory journal* **10**(3), 311–317.
- Duan, W.-X., Hua, R.-X., Yi, W., Shen, L.-J., Jin, Z.-X., Zhao, Y.-H., Yi, D.-H., Chen, W.-S. and Yu, S.-Q. (2012), 'The association between ogg1 ser326cys polymorphism and lung cancer susceptibility: a meta-analysis of 27 studies', *PLoS One* **7**(4), e35970.
- ElaineN.Marieb, K. H. (2012), *Human Anatomy & Physiology*, Pearson.
- Empey, D. W. (1978), 'Diseases of the respiratory system. introduction: structure and function of the lungs.', *British medical journal* **1**(6113), 631.
- Fallahzadeh-Abarghoeei, L., Zahedi, T., Mirabedi, F. and Saadat, M. (2015), 'Allelic prevalence of intron 3 insertion/deletion genetic polymorphism of dna double-strand break repair gene xrccl4 in four healthy iranian populations', *Egyptian Journal of Med-*

ical Human Genetics **16**(3), 215–218.

- Fan, J., Otterlei, M., Wong, H.-K., Tomkinson, A. E. and Wilson III, D. M. (2004), 'Xrcc1 co-localizes and physically interacts with pcna', *Nucleic acids research* **32**(7), 2193–2201.
- Farnebo, L., Stjernström, A., Fredrikson, M., Ansell, A., Garvin, S. and Thunell, L. K. (2015), 'Dna repair genes xpc, xpd, xrcc1, and xrcc3 are associated with risk and survival of squamous cell carcinoma of the head and neck', *DNA repair* **31**, 64–72.
- Faucher, F., Doublié, S. and Jia, Z. (2012), '8-oxoguanine dna glycosylases: one lesion, three subfamilies', *International journal of molecular sciences* **13**(6), 6711–6729.
- Fleiss, J. L. and Gross, A. J. (1991), 'Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique', *Journal of clinical epidemiology* **44**(2), 127–139.
- Fong, K., Sekido, Y., Gazdar, A. and Minna, J. (2003), 'Lung cancer 9: Molecular biology of lung cancer: clinical implications', *Thorax* **58**(10), 892–900.
- Forgacs, E., Zöchbauer-Müller, S., Oláh, E. and Minna, J. D. (2001), 'Molecular genetic abnormalities in the pathogenesis of human lung cancer', *Pathology Oncology Research* **7**(1), 6–13.
- Forget, A. L., Bennett, B. T. and Knight, K. L. (2004), 'Xrcc3 is recruited to dna double strand breaks early and independent of rad51', *Journal of cellular biochemistry* **93**(3), 429–436.
- Fortini, P., Parlanti, E., Sidorkina, O. M., Laval, J. and Dogliotti, E. (1999), 'The type of dna glycosylase determines the base excision repair pathway in mammalian cells', *Journal of Biological Chemistry* **274**(21), 15230–15236.
- Funai, K., Yokose, T., Ishii, G.-i., Araki, K., Yoshida, J., Nishimura, M., Nagai, K., Nishiwaki, Y. and Ochiai, A. (2003), 'Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung', *The American journal of surgical pathology* **27**(7), 978–984.
- Gabel, S. A., DeRose, E. F. and London, R. E. (2013), 'Xrcc1 interaction with the rev1 c-terminal domain suggests a role in post replication repair', *DNA repair* **12**(12), 1105–1113.
- Gangwar, R., Ahirwar, D., Mandhani, A. and Mittal, R. D. (2009), 'Do dna repair genes ogg1, xrcc3 and xrcc7 have an impact on susceptibility to bladder cancer in the north indian population?', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **680**(1), 56–63.
- Gao, Y., Ferguson, D. O., Xie, W., Manis, J. P. *et al.* (2000), 'Interplay of p53 and dna-repair protein xrcc4 in tumorigenesis, genomic stability and development', *Nature* **404**(6780), 897.

- Gasperino, J. (2011), 'Gender is a risk factor for lung cancer', *Medical hypotheses* **76**(3), 328–331.
- Gaughan, E. M., Cryer, S. K., Yeap, B. Y., Jackman, D. M. and Costa, D. B. (2013), 'Family history of lung cancer in never smokers with non-small-cell lung cancer and its association with tumors harboring egfr mutations', *Lung cancer* **79**(3), 193–197.
- Gawełek, E., Drozdowska, B. and Fuchs, A. (2017), 'Radon as a risk factor of lung cancer', *Przegląd epidemiologiczny* **71**(1), 90–98.
- Gelboin, H. V. (1980), 'Benzo [alpha] pyrene metabolism, activation and carcinogenesis: role and regulation of mixed-function oxidases and related enzymes.', *Physiological reviews* **60**(4), 1107–1166.
- Giachino, D. F., Ghio, P., Regazzoni, S., Mandrile, G., Novello, S., Selvaggi, G., Gregori, D., DeMarchi, M. and Scagliotti, G. V. (2007), 'Prospective assessment of xpd lys751gln and xrrc1 arg399gln single nucleotide polymorphisms in lung cancer', *Clinical Cancer Research* **13**(10), 2876–2881.
- Globacon (2012), 'Lung cancer estimated incidence, mortality and prevalence worldwide'.
URL: <http://globocan.iarc.fr/Pages/fact-sheets-cancer.aspx>
- Goldsby, R. E., Lawrence, N. A., Hays, L. E., Olmsted, E. A., Chen, X., Singh, M. and Preston, B. D. (2001), 'Defective dna polymerase-[delta] proofreading causes cancer susceptibility in mice', *Nature medicine* **7**(6), 638–640.
- Goodsell, D. S. (1999), 'The molecular perspective: the ras oncogene', *Stem cells* **17**(4), 235–236.
- Griffiths AJF, Miller JH, S. D. e. a. (2000), *An Introduction to Genetic Analysis. 7th edition.*, W. H. Freeman.
- Gulnaz, A., Sayyed, A. H., Amin, F., Aslam, M. A., Shaikh, R. S., Ali, M. *et al.* (2013), 'Association of xrrc1, xrrc3, and xpd genetic polymorphism with an increased risk of hepatocellular carcinoma because of the hepatitis b and c virus', *European journal of gastroenterology & hepatology* **25**(2), 166–179.
- Guo, Q.-x. and Yang, W.-h. (2013), 'Xrrc1 codon 280 polymorphism and susceptibility to lung cancer: a meta-analysis of the literatures', *Tumor Biology* **34**(5), 2989–2994.
- Guo, S. and Li, X. (2013), 'The relationship between xrrc1 and xrrc3 gene polymorphisms and lung cancer risk in northeastern chinese', *PloS one* **8**(2), e56213.
- Gupta, U., Mir, S. S., Garg, N., Agarwal, S. K., Pande, S. and Mittal, B. (2016), 'Association study of inflammatory genes with rheumatic heart disease in north indian population: A multi-analytical approach', *Immunology letters* **174**, 53–62.
- Gurubhagavatula, S., Liu, G., Park, S., Zhou, W., Su, L., Wain, J. C., Lynch, T. J., Neuberg, D. S. and Christiani, D. C. (2004), 'Xpd and xrrc1 genetic polymorphisms are prognostic factors in advanced nonsmall-cell lung cancer patients treated with platinum

- chemotherapy', *Journal of Clinical Oncology* **22**(13), 2594–2601.
- Hackshaw, A. K., Law, M. R. and Wald, N. J. (1997), 'The accumulated evidence on lung cancer and environmental tobacco smoke', *Bmj* **315**(7114), 980–988.
- Han, J., Zhang, Y. and Li, X. (2014), 'Association between polymorphisms in the *xrcc1* gene and the risk of non-small cell lung cancer.', *Genetics and molecular research: GMR* **14**(4), 12888–12893.
- Han, J., Zhang, Y. and Li, X. (2015), 'Association between polymorphisms in the *xrcc1* gene and the risk of non-small cell lung cancer.', *Genetics and molecular research: GMR* **14**(4), 12888–12893.
- Han Su, Yan Q, Q. H. Z. M. S. J. N. L. H. L. Z. G. M.-H. Q. (2016), 'The *arg194trp* polymorphism in the *xrcc1* gene have no associated with lung cancer risk: A meta-analysis', *Clinics in Oncology* **1**.
- Hanna, J. M. and Onaitis, M. W. (2013), 'Cell of origin of lung cancer', *Journal of carcinogenesis* **12**.
- Hayes, R. B. (1997), 'The carcinogenicity of metals in humans', *Cancer Causes & Control* **8**(3), 371–385.
- He, W., Luo, S., Huang, T., Ren, J., Wu, X., Shao, J. and Zhu, Q. (2012), 'The *ku70-1310c/g* promoter polymorphism is associated with breast cancer susceptibility in chinese han population', *Molecular biology reports* **39**(1), 577–583.
- Hecht, S. S. (1999), 'Tobacco smoke carcinogens and lung cancer', *JNCI: Journal of the National Cancer Institute* **91**(14), 1194–1210.
- Hecht, S. S. (2012), 'Lung carcinogenesis by tobacco smoke', *International journal of cancer* **131**(12), 2724–2732.
- Hegde, M. L., Hegde, P. M., Rao, K. and Mitra, S. (2011), 'Oxidative genome damage and its repair in neurodegenerative diseases: function of transition metals as a double-edged sword', *Journal of Alzheimer's Disease* **24**(supplement2), 183–198.
- Hirata, H., Hinoda, Y., Matsuyama, H., Tanaka, Y., Okayama, N., Suehiro, Y., Zhao, H., Urakami, S., Kawamoto, K., Kawakami, T. *et al.* (2006), 'Polymorphisms of dna repair genes are associated with renal cell carcinoma', *Biochemical and biophysical research communications* **342**(4), 1058–1062.
- Hoffmann, D., Hoffmann, I. and El-Bayoumy, K. (2001), 'The less harmful cigarette: a controversial issue. a tribute to ernst l. wynder', *Chemical research in toxicology* **14**(7), 767–790.
- Hong, C., Xu, Q., Yue, Z., Zhang, Y. and Yuan, Y. (2009), 'Correlation of the sensitivity of np chemotherapy in non-small lung cancer with dna repair gene *xrcc1* polymorphism', *Ai zheng= Aizheng= Chinese journal of cancer* **28**(12), 1291–1297.
- Horikawa, Y., Gu, J. and Wu, X. (2008), 'Genetic susceptibility to bladder cancer with

- an emphasis on gene–gene and gene–environmental interactions’, *Current opinion in urology* **18**(5), 493–498.
- Houtgraaf, J. H., Versmissen, J. and van der Giessen, W. J. (2006), ‘A concise review of dna damage checkpoints and repair in mammalian cells’, *Cardiovascular Revascularization Medicine* **7**(3), 165–172.
- Hsia, T.-C., Chang, W.-S., Chen, W.-C., Liang, S.-J., Tu, C.-Y., Chen, H.-J., Liang, J.-A., Tsai, C.-W., Hsu, C.-M., Tsai, C.-H. *et al.* (2014), ‘Genotype of dna double-strand break repair gene *xrcc7* is associated with lung cancer risk in taiwan males and smokers’, *Anticancer research* **34**(12), 7001–7005.
- Hsia, T.-C., Liu, C.-J., Chu, C.-C., Hang, L.-W., Chang, W.-S., Tsai, C.-W., Wu, C.-I., Lien, C.-S., Liao, W.-L., Ho, C.-Y. *et al.* (2012), ‘Association of dna double-strand break gene *xrcc6* genotypes and lung cancer in taiwan’, *Anticancer research* **32**(3), 1015–1020.
- Hsieh, Y.-H., Chang, W.-S., Tsai, C.-W., Tsai, J.-P., Hsu, C.-M., Jeng, L.-B. and Bau, D.-T. (2015), ‘Dna double-strand break repair gene *xrcc7* genotypes were associated with hepatocellular carcinoma risk in taiwanese males and alcohol drinkers’, *Tumor Biology* **36**(6), 4101–4106.
- Hsu, N.-Y., Wang, H.-C., Wang, C.-H., Chang, C.-L., Chiu, C.-F., Lee, H.-Z., Tsai, C.-W. and Bau, D.-T. (2009), ‘Lung cancer susceptibility and genetic polymorphism of dna repair gene *xrcc4* in taiwan’, *Cancer Biomarkers* **5**(4-5), 159–165.
- Hu, H., Wang, F., Du, X., Zhao, X., Jin, Z. and Hou, M. (2016), ‘Genetic variability of *xrcc1* influences the treatment outcome of gastric cancer’, *Genet Mol Res* **15**.
- Hu, X.-B., Feng, Z., Fan, Y.-C., Xiong, Z.-Y. and Huang, Q.-W. (2011), ‘Polymorphisms in dna repair gene *xrcc1* and increased genetic susceptibility to glioma’, *Asian Pac J Cancer Prev* **12**(11), 2981–2984.
- Hu, Z., Liu, H., Wang, H., Miao, R., Sun, W., Jin, G., Wang, Y., Ma, H., Jin, L., Wei, Q. *et al.* (2008), ‘Tagging single nucleotide polymorphisms in phosphoinositide-3-kinase–related protein kinase genes involved in dna damage checkpoints and lung cancer susceptibility’, *Clinical Cancer Research* **14**(9), 2887–2891.
- Huang, G., Cai, S., Wang, W., Zhang, Q. and Liu, A. (2013), ‘Association between *xrcc1* and *xrcc3* polymorphisms with lung cancer risk: a meta-analysis from case-control studies’, *PloS one* **8**(8), e68457.
- Huang, M., Dinney, C. P., Lin, X., Lin, J., Grossman, H. B. and Wu, X. (2007), ‘High-order interactions among genetic variants in dna base excision repair pathway genes and smoking in bladder cancer susceptibility’, *Cancer Epidemiology and Prevention Biomarkers* **16**(1), 84–91.
- Huang, Z.-H., Hua, D. and Du, X. (2009), ‘Polymorphisms in *p53*, *gstp1* and *xrcc1* predict

- relapse and survival of gastric cancer patients treated with oxaliplatin-based adjuvant chemotherapy', *Cancer chemotherapy and pharmacology* **64**(5), 1001–1007.
- Hung, R. J. and Brennan, P. (2005), 'Large-scale investigation of base excision repair genetic polymorphisms and lung cancer risk in a multicenter study', *Journal of the National Cancer Institute* **97**(8), 567–576.
- Hung, R. J. and Hall, J. (2005), 'Genetic polymorphisms in the base excision repair pathway and cancer risk: a huge review', *American journal of epidemiology* **162**(10), 925–942.
- Improta, G., Sgambato, A., Bianchino, G., Zupa, A., Grieco, V., La Torre, G., Traficante, A. and Cittadini, A. (2008), 'Polymorphisms of the dna repair genes xrccl and xrccl3 and risk of lung and colorectal cancer: a case-control study in a southern italian population', *Anticancer research* **28**(5B), 2941–2946.
- Ito, H., Hamajima, N., Takezaki, T., Matsuo, K., Tajima, K., Hatooka, S., Mitsudomi, T., Suyama, M., Sato, S. and Ueda, R. (2002), 'A limited association of ogg1 ser326cys polymorphism for adenocarcinoma of the lung', *Journal of epidemiology* **12**(3), 258–265.
- Jackson, S. P. (2002), 'Sensing and repairing dna double-strand breaks', *Carcinogenesis* **23**(5), 687–696.
- Jahantigh, D. and Hosseinzadeh Colagar, A. (2017), 'Xrcc5 vntr, xrcc6-61c/g, and xrcc7 6721g/t gene polymorphisms associated with male infertility risk: evidences from case-control and in silico studies', *International journal of endocrinology* **2017**.
- Jahantigh, D., Salimi, S., Mousavi, M., Moossavi, M., Mohammadoo-Khorasani, M., Narooei-nejad, M. and Sandoughi, M. (2015), 'Association between functional polymorphisms of dna double-strand breaks in repair genes xrcc5, xrcc6 and xrcc7 with the risk of systemic lupus erythematosus in south east iran', *DNA and cell biology* **34**(5), 360–366.
- Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E. and Forman, D. (2011), 'Global cancer statistics', *CA: a cancer journal for clinicians* **61**(2), 69–90.
- Jemal, A., Center, M. M., DeSantis, C. and Ward, E. M. (2010), 'Global patterns of cancer incidence and mortality rates and trends', *Cancer Epidemiology and Prevention Biomarkers* **19**(8), 1893–1907.
- Jia, J., Ren, J., Yan, D., Xiao, L. and Sun, R. (2015), 'Association between the xrcc6 polymorphisms and cancer risks: a systematic review and meta-analysis', *Medicine* **94**(1).
- Jiang, H., Lin, Y., Yang, C.-q., Li, Q., Luo, J., Zhou, Y., Xue, J., Wei, W. and Gao, Y. (2013), 'Quantitative assessment of the association between xrcc6 c1310g polymorphism and cancer risk', *Tumor Biology* **34**(2), 779–785.

- Jin, E.-H., Kim, J., Lee, S.-I. and Hong, J. H. (2015), 'Association between polymorphisms in *ape1* and *xrcc1* and the risk of gastric cancer in korean population', *International journal of clinical and experimental medicine* **8**(7), 11484.
- Jin, Z., Zhao, X., Zhang, L., Wang, Y., Yue, W. and Xu, S. (2014), 'Effects of polymorphisms in the *xrcc1*, *xrcc3*, and *xpg* genes on clinical outcomes of platinum-based chemotherapy for treatment of non-small cell lung cancer', *Genet Mol Res* **13**(3), 7617–7625.
- Jindal, S., Malik, S. *et al.* (1988), 'Smoking index-a measure to quantify cumulative smoking exposure', *Lung India* **6**(4), 195.
- Jovanovic, M. and Dynan, W. S. (2006), 'Terminal dna structure and atp influence binding parameters of the dna-dependent protein kinase at an early step prior to dna synapsis', *Nucleic acids research* **34**(4), 1112–1120.
- Kadara, H., Scheet, P., Wistuba, I. I. and Spira, A. E. (2016), 'Early events in the molecular pathogenesis of lung cancer', *Cancer Prevention Research* **9**(7), 518–527.
- Kanaar, R., Hoeijmakers, J. H. and van Gent, D. C. (1998), 'Molecular mechanisms of dna double-strand break repair', *Trends in cell biology* **8**(12), 483–489.
- Kang, S., Ma, Y., Liu, C., Cao, C. *et al.* (2015), 'Association of *xrcc1* gene polymorphisms with risk of non-small cell lung cancer', *International journal of clinical and experimental pathology* **8**(4), 4171.
- Kantor, M., Sobrado, J., Patel, S., Eiseler, S. and Ochner, C. (2017), 'Hereditary colorectal tumors: A literature review on mutyh-associated polyposis', *Gastroenterology research and practice* **2017**.
- Kanwal, M., Ding, X.-J. and Cao, Y. (2017), 'Familial risk for lung cancer', *Oncology letters* **13**(2), 535–542.
- Karahalil, B., Emerce, E., Koçer, B., Han, S., Alkiş, N. and Karakaya, A. (2008), 'The association of *ogg1* ser326cys polymorphism and urinary 8-ohdg levels with lung cancer susceptibility: a hospital-based case-control study in turkey', *Archives of Industrial Hygiene and Toxicology* **59**(4), 241–250.
- Karkucak, M., Yakut, T., Evrensel, T., Deligonul, A., Gulten, T., Ocakoglu, G., Kurt, E., Kanat, O., Cubukcu, E., Sehitoglu, I. *et al.* (2012), 'Xrcc1 gene polymorphisms and risk of lung cancer in turkish patients', *International Journal of Human Genetics* **12**(2), 113–117.
- Karran, P. (2000), 'Dna double strand break repair in mammalian cells', *Current opinion in genetics & development* **10**(2), 144–150.
- Kasahara, M., Osawa, K., Yoshida, K., Miyaishi, A., Osawa, Y., Inoue, N., Tsutou, A., Tabuchi, Y., Tanaka, K., Yamamoto, M. *et al.* (2008), 'Association of mutyh gln324his and apex1 asp148glu with colorectal cancer and smoking in a japanese population',

Journal of Experimental & Clinical Cancer Research **27**(1), 49.

- Kashfi, S. M. H., Golmohammadi, M., Behboudi, F., Nazemalhosseini-Mojarad, E. and Zali, M. R. (2013), 'Mutyh the base excision repair gene family member associated with colorectal cancer polyposis', *Gastroenterology and hepatology from bed to bench* **6**(Suppl 1), S1.
- Katyal, S. and McKinnon, P. J. (2011), 'Disconnecting xrcc1 and dna ligase iii', *Cell Cycle* **10**(14), 2269–2275.
- Ke, H.-G., Li, J., Shen, Y., You, Q.-S., Yan, Y., Dong, H.-X., Liu, J.-H. and Shen, Z.-Y. (2012), 'Prognostic significance of gstp1, xrcc1 and xrcc3 polymorphisms in non-small cell lung cancer patients', *Asian Pacific Journal of Cancer Prevention* **13**(9), 4413–4416.
- Kim, S.-U., Park, S. K., Yoo, K.-Y., Yoon, K.-S., Choi, J. Y., Seo, J.-S., Park, W.-Y., Kim, J.-H., Noh, D.-Y., Ahn, S.-H. *et al.* (2002), 'Xrcc1 genetic polymorphism and breast cancer risk', *Pharmacogenetics and Genomics* **12**(4), 335–338.
- Kim, Y.-J. and M Wilson III, D. (2012), 'Overview of base excision repair biochemistry', *Current molecular pharmacology* **5**(1), 3–13.
- Kinzler KW, V. B. (1997), 'Cancer-susceptibility genes. gatekeepers and caretakers.', *Nature* **386**(6627), 761–763.
- Kiran, M., Saxena, R., Chawla, Y. K. and Kaur, J. (2009), 'Polymorphism of dna repair gene xrcc1 and hepatitis-related hepatocellular carcinoma risk in indian population', *Molecular and cellular biochemistry* **327**(1-2), 7–13.
- Kiran, M., Saxena, R., Kaur, J. *et al.* (2010), 'Distribution of xrcc1 genotypes in north indian population', *Indian Journal of Medical Research* **131**(1), 71.
- Kirschbaum, M. H. and Yarden, Y. (2000), 'The erbb/her family of receptor tyrosine kinases: a potential target for chemoprevention of epithelial neoplasms', *Journal of Cellular Biochemistry* **77**(S34), 52–60.
- Kiyohara, C., Takayama, K. and Nakanishi, Y. (2006), 'Association of genetic polymorphisms in the base excision repair pathway with lung cancer risk: a meta-analysis', *Lung Cancer* **54**(3), 267–283.
- Kiyohara, C., Takayama, K. and Nakanishi, Y. (2010), 'Lung cancer susceptibility and hogg1 ser326cys polymorphism: a meta-analysis', *Cancers* **2**(4), 1813–1829.
- Klinchid, J., Chewaskulyoung, B., Saeteng, S., Lertprasertsuke, N., Kasinrerak, W. and Cressey, R. (2009), 'Effect of combined genetic polymorphisms on lung cancer risk in northern thai women', *Cancer genetics and cytogenetics* **195**(2), 143–149.
- Kohno, T., Kunitoh, H., Mimaki, S., Shiraishi, K., Kuchiba, A., Yamamoto, S. and Yokota, J. (2011), 'Contribution of the tp53, ogg1, chrna3, and hla-dqa1 genes to the risk for lung squamous cell carcinoma', *Journal of Thoracic Oncology* **6**(4), 813–817.

- Kohno, T., Kunitoh, H., Toyama, K., Yamamoto, S., Kuchiba, A., Saito, D., Yanagitani, N., Ishihara, S.-i., Saito, R. and Yokota, J. (2006), 'Association of the ogg1-ser326cys polymorphism with lung adenocarcinoma risk', *Cancer science* **97**(8), 724–728.
- Kondo, T., Yamada, K., Noda, K., Nakayama, H. and Kameda, Y. (2002), 'Radiologic–prognostic correlation in patients with small pulmonary adenocarcinomas', *Lung Cancer* **36**(1), 49–57.
- Kosaka, T., Yatabe, Y., Endoh, H., Kuwano, H., Takahashi, T. and Mitsudomi, T. (2004), 'Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications', *Cancer research* **64**(24), 8919–8923.
- Kraemer, K. H., Patronas, N. J., Schiffmann, R., Brooks, B. P., Tamura, D. and Di-Giovanna, J. J. (2007), 'Xeroderma pigmentosum, trichothiodystrophy and cockayne syndrome: a complex genotype–phenotype relationship', *Neuroscience* **145**(4), 1388–1396.
- Krokan, H. E. and Bjørås, M. (2013), 'Base excision repair', *Cold Spring Harbor perspectives in biology* **5**(4), a012583.
- Kubota, Y., Nash, R. A., Klungland, A., Schär, P., Barnes, D. and Lindahl, T. (1996), 'Reconstitution of dna base excision-repair with purified human proteins: interaction between dna polymerase beta and the xrcc1 protein.', *The EMBO journal* **15**(23), 6662–6670.
- Kumar, A., Pant, M. C., Singh, H. S. and Khandelwal, S. (2011), 'Role of ogg1 ser326cys polymorphism and 8-oxoguanine dna damage in risk assessment of squamous cell carcinoma of head and neck in north indian population', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **726**(2), 227–233.
- Kumar, A., Pant, M. C., Singh, H. S. and Khandelwal, S. (2012), 'Associated risk of xrcc1 and xpd cross talk and life style factors in progression of head and neck cancer in north indian population', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **729**(1), 24–34.
- Ladiges, W. (2006), 'Mouse models of xrcc1 dna repair polymorphisms and cancer', *Oncogene* **25**(11), 1612.
- Lai, C.-Y., Hsieh, L.-L., Tang, R., Santella, R. M., Chang-Chieh, C. R. and Yeh, C.-C. (2016), 'Association between polymorphisms of ape1 and ogg1 and risk of colorectal cancer in taiwan', *World journal of gastroenterology* **22**(12), 3372.
- Lakshmipathy, U. and Campbell, C. (2000), 'Mitochondrial dna ligase iii function is independent of xrcc1', *Nucleic acids research* **28**(20), 3880–3886.
- Lam, T. K., Gallicchio, L., Lindsley, K., Shiels, M., Hammond, E., Tao, X. G., Chen, L., Robinson, K. A., Caulfield, L. E., Herman, J. G. *et al.* (2009), 'Cruciferous vegetable consumption and lung cancer risk: a systematic review', *Cancer Epidemiology and*

- Prevention Biomarkers* **18**(1), 184–195.
- Lamarche, B. J., Orazio, N. I. and Weitzman, M. D. (2010), ‘The mrn complex in double-strand break repair and telomere maintenance’, *FEBS letters* **584**(17), 3682–3695.
- Latimer, K. M. and Mott, T. F. (2015), ‘Lung cancer: diagnosis, treatment principles, and screening.’, *American family physician* **91**(4).
- Le Marchand, L., Donlon, T., Lum-Jones, A., Seifried, A. and Wilkens, L. R. (2002), ‘Association of the hogg1 ser326cys polymorphism with lung cancer risk’, *Cancer Epidemiology and Prevention Biomarkers* **11**(4), 409–412.
- Lee, J.-M., Lee, Y.-C., Yang, S.-Y., Yang, P.-W., Luh, S.-P., Lee, C.-J., Chen, C.-J. and Wu, M.-T. (2001), ‘Genetic polymorphisms of xrcc1 and risk of the esophageal cancer’, *International journal of cancer* **95**(4), 240–246.
- Lengauer, C., Kinzler, K. W. and Vogelstein, B. (1998), ‘Genetic instabilities in human cancers’, *Nature* **396**(6712), 643.
- Letkova, L., Matakova, T., Musak, L., Sarlinova, M., Krutakova, M., Slovakova, P., Kavcova, E., Jakusova, V., Janickova, M., Drgova, A. *et al.* (2013), ‘Dna repair genes polymorphism and lung cancer risk with the emphasis to sex differences’, *Molecular biology reports* **40**(9), 5261–5273.
- Li, D., Liu, H., Jiao, L., Chang, D. Z., Beinart, G., Wolff, R. A., Evans, D. B., Hassan, M. M. and Abbruzzese, J. L. (2006), ‘Significant effect of homologous recombination dna repair gene polymorphisms on pancreatic cancer survival’, *Cancer research* **66**(6), 3323–3330.
- Li, K. and Li, W. (2013), ‘Association between polymorphisms of xrcc1 and adprt genes and ovarian cancer survival with platinum-based chemotherapy in chinese population’, *Molecular and cellular biochemistry* **372**(1-2), 27–33.
- Li, M., Yin, Z., Guan, P., Li, X., Cui, Z., Zhang, J., Bai, W., He, Q. and Zhou, B. (2008), ‘Xrcc1 polymorphisms, cooking oil fume and lung cancer in chinese women nonsmokers’, *Lung cancer* **62**(2), 145–151.
- Li, Q.-w., Lu, C.-r., Ye, M., Xiao, W.-h. and Liang, J. (2012), ‘Evaluation of dna repair gene xrcc1 polymorphism in prediction and prognosis of hepatocellular carcinoma risk’, *Asian Pacific Journal of Cancer Prevention* **13**(1), 191–194.
- Li, R. and Yang, Y. (2011), ‘Genetic polymorphisms in dna double-strand break repair genes xrcc5, xrcc6 and susceptibility to hepatocellular carcinoma’, *Carcinogenesis* **32**(4), 530–536.
- Li, T. and Suo, Q. (2012), ‘Esophageal cancer risk is associated with polymorphisms of dna repair genes msh2 and wrn in chinese population’, *Journal of Thoracic Oncology* **7**(2), 448–452.
- Li, Y., Huang, Y., Cao, Y.-S., Zeng, J., Tong, W.-N., Xu, S.-L. and Zhuo, A.-S. (2013),

- ‘Assessment of the association between *xrcc1* arg399gln polymorphism and lung cancer in chinese’, *Tumor Biology* **34**(6), 3681–3685.
- Li, Z. and Guan, W. (2011), ‘Genetic polymorphism of dna base-excision repair genes (*ape1*, *ogg1* and *xrcc1*) and their correlation with risk of lung cancer in a chinese population’, *Archives of medical research* **42**(3), 226–234.
- Liang, G., Pu, Y. and Yin, L. (2005), ‘Rapid detection of single nucleotide polymorphisms related with lung cancer susceptibility of chinese population’, *Cancer letters* **223**(2), 265–274.
- Liao, W.-Y., Shih, J.-Y., Chang, G.-C., Cheng, Y.-K., Yang, J. C.-H., Chen, Y.-M. and Yu, C.-J. (2012), ‘Genetic polymorphism of *xrcc1* arg399gln is associated with survival in non-small-cell lung cancer patients treated with gemcitabine/platinum’, *Journal of thoracic oncology* **7**(6), 973–981.
- Liddiard, K., Hills, R., Burnett, A. K., Darley, R. L. and Tonks, A. (2010), ‘Ogg1 is a novel prognostic indicator in acute myeloid leukaemia’, *Oncogene* **29**(13), 2005.
- Liu, B., Wei, J., Zou, Z., Qian, X., Nakamura, T., Zhang, W., Ding, Y., Feng, J. and Yu, L. (2007), ‘Polymorphism of *xrcc1* predicts overall survival of gastric cancer patients receiving oxaliplatin-based chemotherapy in chinese population’, *European Journal of Human Genetics* **15**(10), 1049.
- Liu, C. J. and Hsia, T. C. (2010), ‘The joint effect of *hogg1* single nucleotide polymorphism and smoking habit on lung cancer in taiwan’, *Anticancer research* **30**(10), 4141–4145.
- Liu, D., Wu, J., Shi, G., Zhou, H. and Yu, Y. (2014), ‘Role of *xrcc1* and *ercc5* polymorphisms on clinical outcomes in advanced non-small cell lung cancer’, *Genet Mol Res* **13**(2), 3100–3107.
- Liu, L., Wu, C., Wang, Y., Zhong, R., Wang, F., Zhang, X., Duan, S., Lou, J., Yu, D., Tan, W. *et al.* (2010), ‘Association of candidate genetic variations with gastric cardia adenocarcinoma in chinese population: a multiple interaction analysis’, *Carcinogenesis* **32**(3), 336–342.
- Liu, L., Wu, J., Zhong, R., Wu, C., Zou, L., Yang, B., Chen, W., Zhu, B., Duan, S., Yu, D. *et al.* (2013), ‘Multi-loci analysis reveals the importance of genetic variations in sensitivity of platinum-based chemotherapy in non-small-cell lung cancer’, *Molecular carcinogenesis* **52**(12), 923–931.
- Liu, L., Yang, L., Zhang, Y., Xu, Z., Yu, M., Wang, J. and Xiao, Z. (2008), ‘Polymorphisms of *rad51* (g135c) and *xrcc3* (c241t) genes and correlations thereof with prognosis and clinical outcomes of acute myeloid leukemia’, *Zhonghua yi xue za zhi* **88**(6), 378–382.
- Liu, Y. and West, S. C. (2001), ‘Distinct functions of *brca1* and *brca2* in double-strand

- break repair', *Breast Cancer Research* **4**(1), 9.
- Loeb, K. R. and Loeb, L. A. (2000), 'Significance of multiple mutations in cancer', *Carcinogenesis* **21**(3), 379–385.
- Loft, S., Svoboda, P., Kawai, K., Kasai, H., Sørensen, M., Tjønneland, A., Vogel, U., Møller, P., Overvad, K. and Raaschou-Nielsen, O. (2012), 'Association between 8-oxo-7, 8-dihydroguanine excretion and risk of lung cancer in a prospective study', *Free Radical Biology and Medicine* **52**(1), 167–172.
- London, R. E. (2015), 'The structural basis of xrccl-mediated dna repair', *DNA repair* **30**, 90–103.
- López-Cima, M. F., González-Arriaga, P., García-Castro, L., Pascual, T., Marrón, M. G., Puente, X. S. and Tardón, A. (2007), 'Polymorphisms in xpc, xpd, xrccl, and xrccl3 dna repair genes and lung cancer risk in a population of northern Spain', *BMC cancer* **7**(1), 162.
- Lung Cancer* (2012).
- URL:** www.globocan.iarc.fr/Pages/fact-sheets-cancer.aspx
- Lung Cancer* (2017).
- URL:** <https://ghr.nlm.nih.gov/condition/lung-cancer>
- Lung Cancer* (2019).
- URL:** <https://www.cancer.org/cancer/lungcancer.html>
- LUNG CANCER Focus:INDIA* (2007), Technical report.
- Lung cancer statistics* (2015).
- URL:** <https://www.lung-cancer.canceraustralia.gov.au/statistics.html>
- Lung cancer statistics UK* (2015).
- URL:** <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer>
- Lunn, R. M., Langlois, R. G., Hsieh, L. L., Thompson, C. L. and Bell, D. A. (1999), 'Xrccl polymorphisms: effects on aflatoxin b1-dna adducts and glycophorin a variant frequency', *Cancer research* **59**(11), 2557–2561.
- Lynch, H. T., Smyrk, T. C., Watson, P., Lanspa, S. J., Lynch, J. F., Lynch, P. M., Cavalieri, R. J. and Boland, C. R. (1993), 'Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review', *Gastroenterology* **104**(5), 1535–1549.
- MacLeod, M. C. and Tang, M.-s. (1985), 'Interactions of benzo (a) pyrene diol-epoxides with linear and supercoiled dna', *Cancer research* **45**(1), 51–56.
- Maher, R. L., Branagan, A. M. and Morrical, S. W. (2011), 'Coordination of dna replication and recombination activities in the maintenance of genome stability', *Journal of cellular biochemistry* **112**(10), 2672–2682.

- Mahjabeen, I., Baig, R. M., Masood, N., Sabir, M., Inayat, U., Malik, F. A. and Kayani, M. A. (2013), 'Genetic variations in *xrcc1* gene in sporadic head and neck cancer (hnc) patients', *Pathology & Oncology Research* **19**(2), 183–188.
- Malhotra, J., Malvezzi, M., Negri, E., La Vecchia, C. and Boffetta, P. (2016), 'Risk factors for lung cancer worldwide', *European Respiratory Journal* pp. ERJ-00359.
- Malik, P. S. and Raina, V. (2015), 'Lung cancer: Prevalent trends & emerging concepts', *The Indian journal of medical research* **141**(1), 5.
- Mandal, R. K., Gangwar, R., Kapoor, R. and Mittal, R. D. (2012), 'Polymorphisms in base-excision & nucleotide-excision repair genes & prostate cancer risk in north indian population', *The Indian journal of medical research* **135**(1), 64.
- Mandal, R. K., Kapoor, R. and Mittal, R. D. (2010), 'Polymorphic variants of dna repair gene *xrcc3* and *xrcc7* and risk of prostate cancer: a study from north indian population', *DNA and cell biology* **29**(11), 669–674.
- Mandal, R. K., Singh, V., Kapoor, R. and Mittal, R. (2011), 'Do polymorphisms in *xrcc4* influence prostate cancer susceptibility in north indian population?', *Biomarkers* **16**(3), 236–242.
- Manini, P., De Palma, G., Andreoli, R., Marczynski, B., Hanova, M., Mozzoni, P., Naccarati, A., Vodickova, L., Hlavac, P., Mutti, A. *et al.* (2009), 'Biomarkers of nucleic acid oxidation, polymorphism in, and expression of, *hogg1* gene in styrene-exposed workers', *Toxicology letters* **190**(1), 41–47.
- Martin, L. J. (2008), 'Dna damage and repair: relevance to mechanisms of neurodegeneration', *Journal of Neuropathology & Experimental Neurology* **67**(5), 377–387.
- Massion, P. P. and Carbone, D. P. (2003), 'The molecular basis of lung cancer: molecular abnormalities and therapeutic implications', *Respiratory research* **4**(1), 12.
- Matullo, G., Dunning, A., Guarrera, S., Baynes, C., Polidoro, S., Garte, S., Autrup, H., Malaveille, C., Peluso, M., Airoldi, L. *et al.* (2005), 'Dna repair polymorphisms and cancer risk in non-smokers in a cohort study', *Carcinogenesis* **27**(5), 997–1007.
- Matullo, G., Guarrera, S., Carturan, S., Peluso, M., Malaveille, C., Davico, L., Piazza, A. and Vineis, P. (2001), 'Dna repair gene polymorphisms, bulky dna adducts in white blood cells and bladder cancer in a case-control study', *International journal of cancer* **92**(4), 562–567.
- McCarthy, W. J., Meza, R., Jeon, J. and Moolgavkar, S. H. (2012), 'Lung cancer in never smokers: Epidemiology and risk prediction models', *Risk Analysis* **32**(s1).
- Meers, C., Keskin, H. and Storici, F. (2016), 'Dna repair by rna: templated, or not templated, that is the question', *DNA repair* **44**, 17–21.
- Melville, A. and Eastwood, A. (1998), 'Management of lung cancer', *Quality in health care: QHC* **7**(3), 170.

- Metsola, K., Kataja, V., Sillanpää, P., Siivola, P., Heikinheimo, L., Eskelinen, M., Kosma, V.-M., Uusitupa, M. and Hirvonen, A. (2005), 'Xrcc1 and xpd genetic polymorphisms, smoking and breast cancer risk in a finnish case-control study', *Breast Cancer Research* **7**(6), R987.
- Mirsadraee, S., Oswal, D., Alizadeh, Y., Caulo, A. and van Beek, E. J. (2012), 'The 7th lung cancer tnm classification and staging system: Review of the changes and implications', *World journal of radiology* **4**(4), 128.
- Mittal, R. D. and Gangwar, R. (2012), 'Gene variants of xrcc4 and xrcc3 and their association with risk for urothelial bladder cancer', *Molecular biology reports* **39**(2), 1667–1675.
- Mittal, R. D., Mandal, R. K. and Gangwar, R. (2012), 'Base excision repair pathway genes polymorphism in prostate and bladder cancer risk in north indian population', *Mechanisms of ageing and development* **133**(4), 127–132.
- Miyaishi, A., Osawa, K., Osawa, Y., Inoue, N., Yoshida, K., Kasahara, M., Tsutou, A., Tabuchi, Y., Sakamoto, K., Tsubota, N. *et al.* (2009), 'Mutyh gln324his gene polymorphism and genetic susceptibility for lung cancer in a japanese population', *Journal of Experimental & Clinical Cancer Research* **28**(1), 10.
- Mohrenweiser, H. W., Xi, T., Vázquez-Matías, J. and Jones, I. M. (2002), 'Identification of 127 amino acid substitution variants in screening 37 dna repair genes in humans', *Cancer Epidemiology and Prevention Biomarkers* **11**(10), 1054–1064.
- Molina, J. R., Yang, P., Cassivi, S. D., Schild, S. E. and Adjei, A. A. (2008), Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship, *in* 'Mayo Clinic Proceedings', Vol. 83, Elsevier, pp. 584–594.
- Monteiro, E., Varzim, G., Silva, R., Lopes, C. *et al.* (2005), 'Polymorphisms of the human ogg1 gene in laryngeal cancer: implications in radiotherapy response and survival.', *Revue de laryngologie-otologie-rhinologie* **126**(3), 135–140.
- Mucci, L., Wedren, S., Tamimi, R., Trichopoulos, D. and Adami, H.-O. (2001), 'The role of gene–environment interaction in the aetiology of human cancer: examples from cancers of the large bowel, lung and breast', *Journal of internal medicine* **249**(6), 477–493.
- Nagaishi, C. and Nagasawa, N. (1972), *Functional anatomy and histology of the lung*, University Park Press.
- Nagaraju, G. and Scully, R. (2007), 'Minding the gap: the underground functions of brca1 and brca2 at stalled replication forks', *DNA repair* **6**(7), 1018–1031.
- Nakabeppu, Y., Tsuchimoto, D., Furuichi, M. and Sakumi, K. (2004), 'Review the defense mechanisms in mammalian cells against oxidative damage in nucleic acids and their involvement in the suppression of mutagenesis and cell death', *Free radical research* **38**(5), 423–429.

- Nasiri, M., Saadat, I., Omidvari, S. and Saadat, M. (2012), 'Genetic variation in dna repair gene xrcc7 (g6721t) and susceptibility to breast cancer', *Gene* **505**(1), 195–197.
- National Comprehensive Cancer Network (2016).
URL: https://www.nccn.org/professionals/physician_gls/recently-updated.asp
- Natukula, K., Jamil, K., Pingali, U. R., Attili, V. and Madireddy, U. (2013), 'The codon 399 arg/gln xrcc1 polymorphism is associated with lung cancer in indians', *Asian Pac J Cancer Prev* **14**(9), 5275–5279.
- Neumann, A. S., Sturgis, E. M. and Wei, Q. (2005), 'Nucleotide excision repair as a marker for susceptibility to tobacco-related cancers: a review of molecular epidemiological studies', *Molecular carcinogenesis* **42**(2), 65–92.
- New cancer cases in India by 2020 (2015).
URL: <http://icmr.nic.in/icmrsql/archive/2016/7.pdf>
- Nicholson, S. A., Beasley, M. B., Brambilla, E., Hasleton, P. S., Colby, T. V., Sheppard, M. N., Falk, R. and Travis, W. D. (2002), 'Small cell lung carcinoma (sclc): a clinicopathologic study of 100 cases with surgical specimens', *The American journal of surgical pathology* **26**(9), 1184–1197.
- Nohmi, T., Kim, S.-R. and Yamada, M. (2005), 'Modulation of oxidative mutagenesis and carcinogenesis by polymorphic forms of human dna repair enzymes', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **591**(1), 60–73.
- Noronha, V., Pinninti, R., Patil, V. M., Joshi, A. and Prabhash, K. (2016), 'Lung cancer in the indian subcontinent', *South Asian Journal of Cancer* **5**(3), 95.
- Nowacka-Zawisza, M., Wiśnik, E., Wasilewski, A., Skowrońska, M., Forma, E., Bryś, M., Rózański, W. and Krajewska, W. M. (2015), 'Polymorphisms of homologous recombination rad51, rad51b, xrcc2, and xrcc3 genes and the risk of prostate cancer', *Analytical Cellular Pathology* **2015**.
- Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G. *et al.* (2010), 'Three dna polymerases, recruited by different mechanisms, carry out ner repair synthesis in human cells', *Molecular cell* **37**(5), 714–727.
- Oka, S., Leon, J., Tsuchimoto, D., Sakumi, K. and Nakabeppu, Y. (2014), 'Mutyh, an adenine dna glycosylase, mediates p53 tumor suppression via parp-dependent cell death', *Oncogenesis* **3**(10), e121.
- Omiecinski, C. J., Vanden Heuvel, J. P., Perdew, G. H. and Peters, J. M. (2010), 'Xenobiotic metabolism, disposition, and regulation by receptors: from biochemical phenomenon to predictors of major toxicities', *Toxicological Sciences* **120**(suppl_1), S49–S75.
- on the Evaluation of Carcinogenic Risks to Humans, I. W. G., Organization, W. H. and for

- Research on Cancer, I. A. (2004), *Some drinking-water disinfectants and contaminants, including arsenic*, Vol. 84, IARC.
- OpenStax, C. (2014), 'Anatomy & physiology', *Human Anatomy & Physiology*.
- Osann, K. E., Anton-Culver, H., Kurosaki, T. and Taylor, T. (1993), 'Sex differences in lung-cancer risk associated with cigarette smoking', *International journal of cancer* **54**(1), 44–48.
- Osawa, K., Nakarai, C., Uchino, K., Yoshimura, M., Tsubota, N., Takahashi, J. and Kido, Y. (2012), 'Xrcc3 gene polymorphism is associated with survival in japanese lung cancer patients', *International journal of molecular sciences* **13**(12), 16658–16667.
- Ouyang, F.-d., Yang, F.-l., Chen, H.-c., Khan, M. A., Huang, F.-m., Wan, X.-x., Xu, A.-h., Huang, X., Zhou, M.-j., Fang, Q. *et al.* (2013), 'Polymorphisms of dna repair genes xpd, xrcc1, and ogg1, and lung adenocarcinoma susceptibility in chinese population', *Tumor Biology* **34**(5), 2843–2848.
- P Basourakos, S., Li, L., M Aparicio, A., G Corn, P., Kim, J. and C Thompson, T. (2017), 'Combination platinum-based and dna damage response-targeting cancer therapy: evolution and future directions', *Current medicinal chemistry* **24**(15), 1586–1606.
- Pachouri, S. S., Sobti, R. C., Kaur, P. and Singh, J. (2007), 'Contrasting impact of dna repair gene xrcc1 polymorphisms arg399gln and arg194trp on the risk of lung cancer in the north-indian population', *DNA and cell biology* **26**(3), 186–191.
- Panov, S. Z. (2005), 'Molecular biology of the lung cancer', *Radiology and Oncology* **39**(3).
- Papadopoulos, A., Guida, F., Leffondré, K., Cénée, S., Cyr, D., Schmaus, A., Radoi, L., Paget-Bailly, S., Carton, M., Menvielle, G. *et al.* (2014), 'Heavy smoking and lung cancer: are women at higher risk? result of the icare study', *British journal of cancer* **110**(5), 1385.
- Park, J. Y., Lee, S. Y., Jeon, H.-S., Bae, N. C., Chae, S. C., Joo, S., Kim, C. H., Park, J.-H., Kam, S. and Jung, T. H. (2002), 'Polymorphism of the dna repair gene xrcc1 and risk of primary lung cancer', *Cancer Epidemiology and Prevention Biomarkers* **11**(1), 23–27.
- Patil, M. N., Datkhile, K. D., Bhosale, S. J., Chougule, P. G., Vhaval, R. D., Khamkar, T. S., Durgawale, P. P. and Kakade, S. V. (2016), 'Role of genetic polymorphisms in xrcc4, xrcc5, xrcc6 and xrcc7 in breast cancer susceptibility in rural indian population: A hospital based case-control study from maharashtra', *International Journal of Health Sciences and Research (IJHSR)* **6**(11), 24–32.
- Pawelczak, K. S., Andrews, B. J. and Turchi, J. J. (2005), 'Differential activation of dna-pk based on dna strand orientation and sequence bias', *Nucleic acids research* **33**(1), 152–161.
- Peng, Y., Li, Z., Zhang, S., Xiong, Y., Cun, Y., Qian, C., Li, M., Ren, T., Xia, L., Cheng,

- Y. *et al.* (2014), ‘Association of dna base excision repair genes (ogg1, ape1 and xrrc1) polymorphisms with outcome to platinum-based chemotherapy in advanced nonsmall-cell lung cancer patients’, *International journal of cancer* **135**(11), 2687–2696.
- Peto, R., Boreham, J., Lopez, A. D., Thun, M. and Heath, C. (1992), ‘Mortality from tobacco in developed countries: indirect estimation from national vital statistics’, *The Lancet* **339**(8804), 1268–1278.
- Phillips, D. H. (2002), ‘Smoking-related dna and protein adducts in human tissues’, *Carcinogenesis* **23**(12), 1979–2004.
- Pohjola, S. K., Lappi, M., Honkanen, M., Rantanen, L. and Savela, K. (2003), ‘Dna binding of polycyclic aromatic hydrocarbons in a human bronchial epithelial cell line treated with diesel and gasoline particulate extracts and benzo [a] pyrene’, *Mutagenesis* **18**(5), 429–438.
- Poulsen, M. L. M. and Bisgaard, M. (2008), ‘Mutyh associated polyposis (map)’, *Current genomics* **9**(6), 420.
- Prasad, R., Ahuja, R., Singhal, S., Srivastava, A., James, P., Kesarwani, V. and Singh, D. (2010), ‘A case-control study of bidi smoking and bronchogenic carcinoma’, *Annals of thoracic medicine* **5**(4), 238.
- Press, N. A. (1991), *Cells of Origin for Lung Cancer*, National Research Council (US) Panel on Dosimetric Assumptions Affecting the Application of Radon Risk Estimates. **URL:** <https://www.ncbi.nlm.nih.gov/books/NBK234231/>
- Przybylowska, K., Kabzinski, J., Sygut, A., Dziki, L., Dziki, A. and Majsterek, I. (2013), ‘An association selected polymorphisms of xrrc1, ogg1 and mutyh gene and the level of efficiency oxidative dna damage repair with a risk of colorectal cancer’, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **745**, 6–15.
- Pu, X., Ye, Y., Spitz, M. R., Wang, L., Gu, J., Lippman, S. M., Hildebrandt, M. A., Hong, W. K., Minna, J. D., Roth, J. A. *et al.* (2012), ‘Predictors of survival in never-smokers with non-small cell lung cancer: A large-scale, two-phase genetic study’, *Clinical Cancer Research* .
- Qian, B., Zhang, H., Zhang, L., Zhou, X., Yu, H. and Chen, K. (2011), ‘Association of genetic polymorphisms in dna repair pathway genes with non-small cell lung cancer risk’, *Lung Cancer* **73**(2), 138–146.
- Qiu, M., Xu, L., Yang, X., Ding, X., Hu, J., Jiang, F., Xu, L. and Yin, R. (2013), ‘Xrrc3 thr241met is associated with response to platinum-based chemotherapy but not survival in advanced non-small cell lung cancer’, *PloS one* **8**(10), e77005.
- RadiologyInfo.org (2017), ‘Lung cancer treatment’.
- Rai, R., Kim, J. J., Misra, S., Kumar, A. and Mittal, B. (2015), ‘A multiple interaction analysis reveals adrb3 as a potential candidate for gallbladder cancer predisposition

- via a complex interaction with other candidate gene variations', *International journal of molecular sciences* **16**(12), 28038–28049.
- Ramachandran, S., Ramadas, K., Hariharan, R., Kumar, R. R. and Pillai, M. R. (2006), 'Single nucleotide polymorphisms of dna repair genes xrccl and xpd and its molecular mapping in indian oral cancer', *Oral oncology* **42**(4), 350–362.
- Rami-Porta, R., Crowley, J. J. and Goldstraw, P. (2009), 'Review the revised tnm staging system for lung cancer', *Ann Thorac Cardiovasc Surg* **15**(1), 5.
- Rastogi, R. P., Kumar, A., Tyagi, M. B., Sinha, R. P. *et al.* (2010), 'Molecular mechanisms of ultraviolet radiation-induced dna damage and repair', *Journal of nucleic acids* **2010**.
- Ratnasinghe, D., Yao, S.-X., Tangrea, J. A., Qiao, Y.-L., Andersen, M. R., Barrett, M. J., Giffen, C. A., Erozan, Y., Tockman, M. S. and Taylor, P. R. (2001), 'Polymorphisms of the dna repair gene xrccl and lung cancer risk', *Cancer Epidemiology and Prevention Biomarkers* **10**(2), 119–123.
- R.B. Ji, Y.S. Qian, A. H. and Hu, Y. (2015), 'Dna repair gene xrccl t241m polymorphism and susceptibility to hepatocellular carcinoma in a chinese population: a meta-analysis', *Genetics and Molecular Research* **14** (4):, 15988–15995.
- Reisz, J. A., Bansal, N., Qian, J., Zhao, W. and Furdul, C. M. (2014), 'Effects of ionizing radiation on biological moleculesmechanisms of damage and emerging methods of detection', *Antioxidants & redox signaling* **21**(2), 260–292.
- Ren, S., Zhou, S., Wu, F., Zhang, L., Li, X., Zhang, J., Xu, J., Lv, M. and Zhou, C. (2012), 'Association between polymorphisms of dna repair genes and survival of advanced nsccl patients treated with platinum-based chemotherapy', *lung cancer* **75**(1), 102–109.
- Riely, G. J., Kris, M. G., Rosenbaum, D., Marks, J., Li, A., Chitale, D. A., Nafa, K., Riedel, E. R., Hsu, M., Pao, W. *et al.* (2008), 'Frequency and distinctive spectrum of kras mutations in never smokers with lung adenocarcinoma', *Clinical Cancer Research* **14**(18), 5731–5734.
- Risch, H. A., Howe, G. R., Jain, M., Burch, J. D., Holowaty, E. J. and Miller, A. B. (1993), 'Are female smokers at higher risk for lung cancer than male smokers? a case-control analysis by histologic type', *American journal of epidemiology* **138**(5), 281–293.
- Ritchie, M. D., Hahn, L. W., Roodi, N., Bailey, L. R., Dupont, W. D., Parl, F. F. and Moore, J. H. (2001), 'Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer', *The American Journal of Human Genetics* **69**(1), 138–147.
- Ritchie, M. D. and Motsinger, A. A. (2005), 'Multifactor dimensionality reduction for detecting gene–gene and gene–environment interactions in pharmacogenomics studies'.
- Rivera-Calzada, A., Spagnolo, L., Pearl, L. H. and Llorca, O. (2007), 'Structural model of full-length human ku70–ku80 heterodimer and its recognition of dna and dna-pkcs',

- EMBO reports* **8**(1), 56–62.
- Rivera, M. P. and Stover, D. E. (2004), ‘Gender and lung cancer’, *Clinics in chest medicine* **25**(2), 391–400.
- Rivlin, N., Brosh, R., Oren, M. and Rotter, V. (2011), ‘Mutations in the p53 tumor suppressor gene: important milestones at the various steps of tumorigenesis’, *Genes & cancer* **2**(4), 466–474.
- Rom, W. N., Hay, J. G., Lee, T. C., Jiang, Y. and Tchou-Wong, K.-M. (2000), ‘Molecular and genetic aspects of lung cancer’, *American journal of respiratory and critical care medicine* **161**(4), 1355–1367.
- Roos, W. P. and Kaina, B. (2006), ‘Dna damage-induced cell death by apoptosis’, *Trends in molecular medicine* **12**(9), 440–450.
- Rosell, R., Lord, R. V., Taron, M. and Reguart, N. (2002), ‘Dna repair and cisplatin resistance in non-small-cell lung cancer’, *Lung cancer* **38**(3), 217–227.
- Rota, M., Bosetti, C., Boccia, S., Boffetta, P. and La Vecchia, C. (2014), ‘Occupational exposures to polycyclic aromatic hydrocarbons and respiratory and urinary tract cancers: an updated systematic review and a meta-analysis to 2014’, *Archives of toxicology* **88**(8), 1479–1490.
- Rubin, B., Skarin, A., Pisick, E., Rizk, M. and Salgia, R. (2001), ‘Use of cytokeratins 7 and 20 in determining the origin of metastatic carcinoma of unknown primary, with special emphasis on lung cancer’, *European Journal of Cancer Prevention* **10**(1), 77–82.
- Rusch, V. W. (2015), *LUNG CANCER WORKUP AND STAGING*, Elsevier.
- Saadat, M. and Saadat, I. (2012), ‘Prevalence of g6721t polymorphism of xrcc7 in an iranian population’, *EXCLI journal* **11**, 93.
- Saadat, M. and Saadat, S. (2015), ‘Susceptibility to breast cancer and intron 3 ins/del genetic polymorphism of dna double-strand break repair gene xrcc4’, *Journal of medical biochemistry* **34**(4), 409–413.
- Sacerdote, C., Guarrera, S., Ricceri, F., Pardini, B., Polidoro, S., Allione, A., Critelli, R., Russo, A., Andrew, A. S., Ye, Y. *et al.* (2013), ‘Polymorphisms in the xrcc1 gene modify survival of bladder cancer patients treated with chemotherapy’, *International journal of cancer* **133**(8), 2004–2009.
- Saikia, B. J., Phukan, R. K., Sharma, S. K., Sekhon, G. S. and Mahanta, J. (2014), ‘Interaction of xrcc1 and xpd gene polymorphisms with lifestyle and environmental factors regarding susceptibility to lung cancer in a high incidence population in north east india’, *Asian Pac J Cancer Prev* **15**(5), 1993–9.
- Saito, S., Kurosawa, A. and Adachi, N. (2016), ‘Mutations in xrcc4 cause primordial dwarfism without causing immunodeficiency’, *Journal of human genetics* **61**(8), 679.
- Sak, S. C., Barrett, J. H., Paul, A. B., Bishop, D. T. and Kiltie, A. E. (2007), ‘Dna repair

- gene *xrcc1* polymorphisms and bladder cancer risk', *BMC genetics* **8**(1), 13.
- Sameer, A. S., Nissar, S., Abdullah, S., Chowdri, N. A. and Siddiqi, M. A. (2012), 'Dna repair gene 8-oxoguanine dna glycosylase ser326cys polymorphism and colorectal cancer risk in a kashmiri population', *DNA and cell biology* **31**(4), 541–546.
- Samet, J. M., Avila-Tang, E., Boffetta, P., Hannan, L. M., Olivo-Marston, S., Thun, M. J. and Rudin, C. M. (2009), 'Lung cancer in never smokers: clinical epidemiology and environmental risk factors', *Clinical Cancer Research* **15**(18), 5626–5645.
- Sancar, A. (1995), 'Dna repair in humans', *Annual review of genetics* **29**(1), 69–105.
- Santos, P. A. R. d., Rocha, R. S. d., Pipkin, M., Silveira, M. L. d., Cypel, M., Rios, J. O. and Pinto, J. A. L. d. F. (2007), 'Concordance between clinical and pathological staging in patients with stages i or ii non-small cell lung cancer subjected to surgical treatment', *Jornal Brasileiro de Pneumologia* **33**(6), 647–654.
- Sasco, A., Secretan, M. and Straif, K. (2004), 'Tobacco smoking and cancer: a brief review of recent epidemiological evidence', *Lung cancer* **45**, S3–S9.
- Sato, M., Shames, D. S., Gazdar, A. F. and Minna, J. D. (2007), 'A translational view of the molecular pathogenesis of lung cancer', *Journal of thoracic oncology* **2**(4), 327–343.
- Schiller, J. H., Harrington, D., Belani, C. P., Langer, C., Sandler, A., Krook, J., Zhu, J. and Johnson, D. H. (2002), 'Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer', *New England Journal of Medicine* **346**(2), 92–98.
- Schneider, J., Classen, V., Bernges, U. and Philipp, M. (2005), 'Xrcc1 polymorphism and lung cancer risk in relation to tobacco smoking', *International journal of molecular medicine* **16**(4), 709–716.
- Sekido, Y., Fong, K. M. and Minna, J. D. (2003), 'Molecular genetics of lung cancer', *Annual review of medicine* **54**(1), 73–87.
- Sellers, T. A., Potter, J. D., Bailey-Wilson, J. E., Rich, S. S., Rothschild, H. and Elston, R. C. (1992), 'Lung cancer detection and prevention: evidence for an interaction between smoking and genetic predisposition', *Cancer research* **52**(9 Supplement), 2694s–2697s.
- Shellard, S., Fichtinger-Schepman, A., Lazo, J. S. and Hill, B. T. (1993), 'Evidence of differential cisplatin-dna adduct formation, removal and tolerance of dna damage in three human lung carcinoma cell lines.', *Anti-cancer drugs* **4**(4), 491–500.
- SHEN, M., BERNDT, S. I., ROTHMAN, N., MUMFORD, J. L., HE, X., YEAGER, M., WELCH, R., CHANOCK, S., KEOHAVONG, P., DONAHUE, M. *et al.* (2005), 'Polymorphisms in the dna base excision repair genes *apex1* and *xrcc1* and lung cancer risk in xuan wei, china', *Anticancer Research* **25**(1B), 537–542.
- Shen, M. R., Jones, I. M. and Mohrenweiser, H. (1998), 'Nonconservative amino acid substitution variants exist at polymorphic frequency in dna repair genes in healthy

- humans', *Cancer research* **58**(4), 604–608.
- Shen, X.-y., Lu, F.-z., Wu, Y., Zhao, L.-t. and Lin, Z.-f. (2013), 'Xrcc3 thr241met polymorphism and clinical outcomes of nslc patients receiving platinum-based chemotherapy: a systematic review and meta-analysis', *PloS one* **8**(8), e69553.
- Shihadeh, A., Schubert, J., Klaiany, J., El Sabban, M., Luch, A. and Saliba, N. A. (2015), 'Toxicant content, physical properties and biological activity of waterpipe tobacco smoke and its tobacco-free alternatives', *Tobacco control* pp. tobaccocontrol–2014.
- Shinmura, K., Goto, M., Suzuki, M., Tao, H., Yamada, H., Igarashi, H., Matsuura, S., Maeda, M., Konno, H., Matsuda, T. *et al.* (2011), 'Reduced expression of mutyh with suppressive activity against mutations caused by 8-hydroxyguanine is a novel predictor of a poor prognosis in human gastric cancer', *The Journal of pathology* **225**(3), 414–423.
- ShollLM (2016), 'The molecular pathology of lung cancer.', *ELSEVIER* **9**(3), 353–78.
- Shrivastav, M., De Haro, L. P. and Nickoloff, J. A. (2008), 'Regulation of dna double-strand break repair pathway choice', *Cell research* **18**(1), 134.
- Simonelli, V., Narciso, L., Dogliotti, E. and Fortini, P. (2005), 'Base excision repair intermediates are mutagenic in mammalian cells', *Nucleic acids research* **33**(14), 4404–4411.
- Sipley, J. D., Menninger, J. C., Hartley, K. O., Ward, D. C., Jackson, S. P. and Anderson, C. W. (1995), 'Gene for the catalytic subunit of the human dna-activated protein kinase maps to the site of the xrcc7 gene on chromosome 8', *Proceedings of the National Academy of Sciences* **92**(16), 7515–7519.
- Sliwinski, T., Przybyłowska, K., Markiewicz, L., Rusin, P., Pietruszewska, W., Zelinska-Blizniewska, H., Olszewski, J., Morawiec-Sztandera, A., Mlynarski, W. and Majsterek, I. (2011), 'Mutyh tyr165cys, ogg1 ser326cys and xpd lys751gln polymorphisms and head neck cancer susceptibility: a case control study', *Molecular biology reports* **38**(2), 1251–1261.
- Smith, T. R., Levine, E. A., Perrier, N. D., Miller, M. S., Freimanis, R. I., Lohman, K., Case, L. D., Xu, J., Mohrenweiser, H. W. and Hu, J. J. (2003), 'Dna-repair genetic polymorphisms and breast cancer risk', *Cancer Epidemiology and Prevention Biomarkers* **12**(11), 1200–1204.
- Spiro, S. G. and Silvestri, G. A. (2005), 'The treatment of advanced non-small cell lung cancer', *Current opinion in pulmonary medicine* **11**(4), 287–291.
- Spitz, M. R., Wei, Q., Dong, Q., Amos, C. I. and Wu, X. (2003), 'Genetic susceptibility to lung cancer: the role of dna damage and repair', *Cancer Epidemiology and Prevention Biomarkers* **12**(8), 689–698.
- Sreeja, L., Syamala, V. S., Syamala, V., Hariharan, S., Raveendran, P. B., Vijayalekshmi,

- R., Madhavan, J. and Ankathil, R. (2008), 'Prognostic importance of dna repair gene polymorphisms of xrccl arg399gln and xpd lys751gln in lung cancer patients from india', *Journal of cancer research and clinical oncology* **134**(6), 645–652.
- Srivastava, A., Srivastava, K., Pandey, S. N., Choudhuri, G. and Mittal, B. (2009), 'Single-nucleotide polymorphisms of dna repair genes ogg1 and xrccl: association with gallbladder cancer in north indian population', *Annals of surgical oncology* **16**(6), 1695–1703.
- Srivastava, D. K., Husain, I., Arteaga, C. L. and Wilson, S. H. (1999), 'Dna polymerase β expression differences in selected human tumors and cell lines', *Carcinogenesis* **20**(6), 1049–1054.
- Srivastava, K., Srivastava, A. and Mittal, B. (2010), 'Polymorphisms in ercc2, msh2, and ogg1 dna repair genes and gallbladder cancer risk in a population of northern india', *Cancer* **116**(13), 3160–3169.
- Staging (2018).
URL: <https://www.ncbi.nlm.nih.gov/books/NBK234231/>
- Steenland, K., Mannetje, A., Boffetta, P., Stayner, L., Attfield, M., Chen, J., Dosemeci, M., DeKlerk, N., Hnizdo, E., Koskela, R. *et al.* (2001), 'Pooled exposure–response analyses and risk assessment for lung cancer in 10 cohorts of silica-exposed workers: an iarc multicentre study', *Cancer Causes & Control* **12**(9), 773–784.
- Stern, M. C., Umbach, D. M., van Gils, C. H., Lunn, R. M. and Taylor, J. A. (2001), 'Dna repair gene xrccl polymorphisms, smoking, and bladder cancer risk', *Cancer Epidemiology and Prevention Biomarkers* **10**(2), 125–131.
- Sterpone, S. and Cozzi, R. (2010), 'Influence of xrccl genetic polymorphisms on ionizing radiation-induced dna damage and repair', *Journal of nucleic acids* **2010**.
- Stoehlmacher, J., Ghaderi, V., Iobal, S., Groshen, S., Tsao-Wei, D., Park, D. and Lenz, H.-J. (2001), 'A polymorphism of the xrccl gene predicts for response to platinum based treatment in advanced colorectal cancer.', *Anticancer research* **21**(4B), 3075–3079.
- Su, Y., Zhang, H., Xu, F., Kong, J., Yu, H. and Qian, B. (2015), 'Dna repair gene polymorphisms in relation to non-small cell lung cancer survival', *Cellular Physiology and Biochemistry* **36**(4), 1419–1429.
- Subramanian, J. and Govindan, R. (2008), 'Molecular genetics of lung cancer in people who have never smoked', *The lancet oncology* **9**(7), 676–682.
- Syamala, V. S., Syamala, V., Sreedharan, H., Raveendran, P. B., Kuttan, R. and Ankathil, R. (2009), 'Contribution of xpd (lys751gln) and xrccl (arg399gln) polymorphisms in familial and sporadic breast cancer predisposition and survival: an indian report', *Pathology & Oncology Research* **15**(3), 389.
- Synowiec, E., Blasiak, J., Zaras, M., Szaflik, J. and Szaflik, J. P. (2012), 'Association

- between polymorphisms of the dna base excision repair genes mutyh and hogg1 and age-related macular degeneration', *Experimental eye research* **98**, 58–66.
- Szaflik, J. P., Cuchra, M., Przybylowska-Sygut, K., Dziki, L., Kurowska, A. K., Gacek, M., Drzewoski, J., Szaflik, J. and Majsterek, I. (2013), 'Association of the 399arg/gln xrcc1, the 194 arg/trp xrcc1, the 326ser/cys ogg1, and the 324gln/his mutyh gene polymorphisms with clinical parameters and the risk for development of primary open-angle glaucoma', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **753**(1), 12–22.
- Takanami, T., Nakamura, J., Kubota, Y. and Horiuchi, S. (2005), 'The arg280his polymorphism in x-ray repair cross-complementing gene 1 impairs dna repair ability', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **582**(1), 135–145.
- Takei, H., Asamura, H., Maeshima, A., Suzuki, K., Kondo, H., Niki, T., Yamada, T., Tsuchiya, R. and Matsuno, Y. (2002), 'Large cell neuroendocrine carcinoma of the lung: a clinicopathologic study of eighty-seven cases', *The Journal of thoracic and cardiovascular surgery* **124**(2), 285–292.
- Takiguchi, Y., Kurimasa, A., Chen, F., Pardington, P. E., Kyama, T., Okinaka, R. T., Moyzis, R. and Chen, D. J. (1996), 'Genomic structure and chromosomal assignment of the mouse ku70 gene', *Genomics* **35**(1), 129–135.
- Tamura, T., Kurishima, K., Nakazawa, K., Kagohashi, K., Ishikawa, H., Satoh, H. and Hizawa, N. (2015), 'Specific organ metastases and survival in metastatic non-small-cell lung cancer', *Molecular and clinical oncology* **3**(1), 217–221.
- Tansavatdi, K., McClain, B. and Herrington, D. (2004), 'The effects of smoking on estradiol metabolism.', *Minerva ginecologica* **56**(1), 105–114.
- Tas, F., Ciftci, R., Kilic, L. and Karabulut, S. (2013), 'Age is a prognostic factor affecting survival in lung cancer patients', *Oncology letters* **6**(5), 1507–1513.
- Taxak, N. and Bharatam, P. V. (2014), 'Drug metabolism a fascinating link between chemistry and biology'.
- Taylor, R. M., Thistlethwaite, A. and Caldecott, K. W. (2002), 'Central role for the xrcc1 brct i domain in mammalian dna single-strand break repair', *Molecular and cellular biology* **22**(8), 2556–2563.
- TerryDesJardins (2008), *Cardiopulmonary Anatomy Physiology Essentials of Respiratory Care*, Thomson Delmar Learning.
- Thomas, A., Rajan, A. and Giaccone, G. (2012), 'Tyrosine kinase inhibitors in lung cancer', *Hematology/Oncology Clinics* **26**(3), 589–605.
- Torgovnick, A. and Schumacher, B. (2015), 'Dna repair mechanisms in cancer development and therapy', *Frontiers in genetics* **6**, 157.
- Travis, W. D. and Harris, C. (2004), *Pathology and genetics of tumours of the lung*,

- pleura, thymus and heart*, Feance: IARC Press, 2004.
- Tseng, R.-C., Hsieh, F.-J., Shih, C.-M., Hsu, H.-S., Chen, C.-Y. and Wang, Y.-C. (2009), 'Lung cancer susceptibility and prognosis associated with polymorphisms in the non-homologous end-joining pathway genes', *Cancer* **115**(13), 2939–2948.
- Tubbs, A. and Nussenzweig, A. (2017), 'Endogenous dna damage as a source of genomic instability in cancer', *Cell* **168**(4), 644–656.
- Tu.J:Inthavong, K.:Ahmadi, G. (2013), *Computational Fluid and Particle Dynamics in the Human Resoiratory System*.
- Type of Tumors (2016).
URL: <http://pathology.jhu.edu/pc/BasicTypes1.php?area=ba>
- Upadhyay, R., Malik, M. A., Zargar, S. A. and Mittal, B. (2010), 'Ogg1 ser326cys polymorphism and susceptibility to esophageal cancer in low and high at-risk populations of northern india', *Journal of gastrointestinal cancer* **41**(2), 110–115.
- Uppal, V., Mehndiratta, M., Mohapatra, D., Grover, R. K. and Puri, D. (2014), 'Xrcc-1 gene polymorphism (arg399gln) and susceptibility to development of lung cancer in cohort of north indian population: A pilot study', *Journal of clinical and diagnostic research: JCDR* **8**(11), CC17.
- UyBico, S. J., Wu, C. C., Suh, R. D., Le, N. H., Brown, K. and Krishnam, M. S. (2010), 'Lung cancer staging essentials: The new tnm staging system and potential imaging pitfalls 1', *Radiographics* **30**(5), 1163–1181.
- van der Kemp, P. A., Charbonnier, J.-B., Audebert, M. and Boiteux, S. (2004), 'Catalytic and dna-binding properties of the human ogg1 dna n-glycosylase/ap lyase: biochemical exploration of h270, q315 and f319, three amino acids of the 8-oxoguanine-binding pocket', *Nucleic acids research* **32**(2), 570–578.
- Varga, T. and Aplan, P. D. (2005), 'Chromosomal aberrations induced by double strand dna breaks', *DNA repair* **4**(9), 1038–1046.
- Vineis, P., Manuguerra, M., Kavvoura, F. K., Guarrera, S., Allione, A., Rosa, F., Di Gregorio, A., Polidoro, S., Saletta, F., Ioannidis, J. P. *et al.* (2009), 'A field synopsis on low-penetrance variants in dna repair genes and cancer susceptibility', *Journal of the National Cancer Institute* **101**(1), 24–36.
- Wang, G. and Wang, S. (2009), 'Polymorphisms in xrcc5, xrcc6, xrcc7 genes are involved in dna double-strand breaks (dsbs) repair associated with the risk of acute myeloid leukemia (aml) in chinese population', *Journal of Nanjing Medical University* **23**(2), 93–99.
- Wang, H. and Liu, Z. (2014), 'Polymorphism of dna repair gene xrcc1 and lung cancer risk', *Cell biochemistry and biophysics* **70**(3), 1881–1886.
- Wang, J.-Y. and Cai, Y. (2014), 'X-ray repair cross-complementing group 1 codon

- 399 polymorphism and lung cancer risk: an updated meta-analysis', *Tumor Biology* **35**(1), 411–418.
- Wang, J., Zhao, Y., Jiang, J., Gajalakshmi, V., Kuriki, K., Nakamura, S., Akasaka, S., Ishikawa, H., Suzuki, S., Nagaya, T. *et al.* (2010), 'Polymorphisms in dna repair genes *xrcc1*, *xrcc3* and *xpd*, and colorectal cancer risk: a case-control study in an indian population', *Journal of cancer research and clinical oncology* **136**(10), 1517–1525.
- Wang, L.-E., Bondy, M. L., Shen, H., El-Zein, R., Aldape, K., Cao, Y., Pudavalli, V., Levin, V. A., Yung, W. A. and Wei, Q. (2004), 'Polymorphisms of dna repair genes and risk of glioma', *Cancer research* **64**(16), 5560–5563.
- Wang, S.-Y., Peng, L., Li, C.-P., Li, A.-P., Zhou, J.-W., Zhang, Z.-D. and Liu, Q.-Z. (2008), 'Genetic variants of the *xrcc7* gene involved in dna repair and risk of human bladder cancer', *International journal of urology* **15**(6), 534–539.
- Wang, X., Zhang, K., Liu, X., Liu, B. and Wang, Z. (2015), 'Association between *xrcc1* and *xrcc3* gene polymorphisms and risk of thyroid cancer', *International journal of clinical and experimental pathology* **8**(3), 3160.
- Wang, Y. and Yang, H. (2009), 'Association between x-ray repair cross complementing group 1 codon 399 and 194 polymorphisms and lung cancer risk: a meta-analysis', *Cancer letters* **285**(2), 134–140.
- Waters, T. R., Gallinari, P., Jiricny, J. and Swann, P. F. (1999), 'Human thymine dna glycosylase binds to apurinic sites in dna but is displaced by human apurinic endonuclease 1', *Journal of Biological Chemistry* **274**(1), 67–74.
- Weaver, D. A., Crawford, E. L., Warner, K. A., Elkhairi, F., Khuder, S. A. and Willey, J. C. (2005), 'Abcc5, *ercc2*, *xpa* and *xrcc1* transcript abundance levels correlate with cisplatin chemoresistance in non-small cell lung cancer cell lines', *Molecular cancer* **4**(1), 18.
- Wei, Q., Cheng, L., Amos, C. I., Wang, L.-E., Guo, Z., Hong, W. K. and Spitz, M. R. (2000), 'Repair of tobacco carcinogen-induced dna adducts and lung cancer risk: a molecular epidemiologic study', *Journal of the National Cancer Institute* **92**(21), 1764–1772.
- Wei, Q., Cheng, L., Hong, W. K. and Spitz, M. R. (1996), 'Reduced dna repair capacity in lung cancer patients', *Cancer research* **56**(18), 4103–4107.
- Weiss (2005), 'An introduction to lung cancer'.
- URL:** <http://cancergrace.org/lung/2010/04/05/an-introduction-to-lung-cancer/>
- Weissferdt, A., . M. C. A. (2012), *Diagnostic Pathology of Pleuropulmonary Neoplasia*, Springer Science & Business Media.
- WHO (2018).
- URL:** <http://www.who.int/mediacentre/factsheets/fs297/en/>

- Wiencke, J. K., Kelsey, K. T., Varkonyi, A., Semey, K., Wain, J. C., Mark, E. and Christiani, D. C. (1995), 'Correlation of dna adducts in blood mononuclear cells with tobacco carcinogen-induced damage in human lung', *Cancer research* **55**(21), 4910–4914.
- Williams, G. T., Critchlow, M. R., Hedge, V. L. and O'Hare, K. B. (1998), 'Molecular failure of apoptosis: inappropriate cell survival and mutagenesis?', *Toxicology letters* **102**, 485–489.
- Wistuba, I. I., Mao, L. and Gazdar, A. F. (2002), 'Smoking molecular damage in bronchial epithelium', *Oncogene* **21**(48), 7298.
- Wu, J., Chen, Y., Wang, L., Yang, Y., Deng, C., Hou, B., He, Z. and Chen, J. (2014), 'Implication of polymorphisms in dna repair genes with an increased risk of hepatocellular carcinoma', *Genet Mol Res* **13**(2), 3812–3818.
- Xu, Z., Hua, L.-X., Qian, L.-X., Yang, J., Wang, X.-R., Zhang, W. and Wu, H.-F. (2007), 'Relationship between xrccl polymorphisms and susceptibility to prostate cancer in men from han, southern china', *Asian journal of andrology* **9**(3), 331–338.
- Yan, L.-J. (2014), 'Positive oxidative stress in aging and aging-related disease tolerance', *Redox Biology* **2**, 165–169.
- Yancik, R. (1997), 'Cancer burden in the aged', *Cancer* **80**(7), 1273–1283.
- Yang, H., Shao, F., Wang, H. and Wang, Y. (2015), 'Note of clarification of data in the paper entitled no association between xrccl gene arg194trp polymorphism and risk of lung cancer: evidence based on an updated cumulative meta-analysis'.
- Yang, I. A., Holloway, J. W. and Fong, K. M. (2013), 'Genetic susceptibility to lung cancer and co-morbidities', *Journal of thoracic disease* **5**(Suppl 5), S454.
- Yang, M.-D., Wang, H.-C., Chang, W.-S., Tsai, C.-W. and Bau, D.-T. (2011), 'Genetic polymorphisms of dna double strand break gene ku70 and gastric cancer in taiwan', *BMC cancer* **11**(1), 174.
- Yao, C.-Y., Huang, X.-E., Li, C., Shen, H.-B., Shi, M.-Q., Feng, J.-F., Pan, L.-X. and Tang, J.-H. (2009), 'Lack of influence of xrccl and xpd gene polymorphisms on outcome of platinum-based chemotherapy for advanced non small cell lung cancers', *Asian Pac J Cancer Prev* **10**(5), 859–64.
- Yin, J., Qi, R., Ma, Y., Sun, Z. and Wang, H. (2007), 'Genetic polymorphisms of dna repair gene: Xrccl pro206pro and gln632gln in a chinese han population', *Biochemical genetics* **45**(11-12), 815–821.
- Yin, J. and Vogel, U. (2007), 'The dna repair gene xrccl and genetic susceptibility of lung cancer in a northeastern chinese population', *Lung Cancer* **56**(2), 153–160.
- Yin, J., Vogel, U., Gerdes, L. U., Dybdahl, M., Bolund, L. and Nexø, B. A. (2003), 'Twelve single nucleotide polymorphisms on chromosome 19q13. 2-13.3: linkage dise-

- quilibria and associations with basal cell carcinoma in danish psoriatic patients', *Biochemical genetics* **41**(1-2), 27–37.
- Yin, J. and Wang, C. (2012), 'No evidence of association between the synonymous polymorphisms in *xrcc1* and *ercc2* and breast cancer susceptibility among nonsmoking chinese', *Gene* **503**(1), 118–122.
- Yin, M. and Liao, Z. (2012), 'Genetic variants of the nonhomologous end joining gene *lig4* and severe radiation pneumonitis in nonsmall cell lung cancer patients treated with definitive radiotherapy', *Cancer* **118**(2), 528–535.
- Yin, Z., Zhou, B., He, Q., Li, M., Guan, P., Li, X., Cui, Z., Xue, X., Su, M., Ma, R. *et al.* (2009), 'Association between polymorphisms in dna repair genes and survival of non-smoking female patients with lung adenocarcinoma', *BMC cancer* **9**(1), 439.
- Yong, Y. and Lin, H. (2005), 'Shesis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci', *Cell research* **15**(2), 97.
- Yoo, S. S., Jin, C., Jung, D. K., Choi, Y. Y., Choi, J. E., Lee, W. K., Lee, S. Y., Lee, J., Cha, S. I., Kim, C. H. *et al.* (2015), 'Putative functional variants of *xrcc1* identified by regulomedb were not associated with lung cancer risk in a korean population', *Cancer genetics* **208**(1), 19–24.
- Yoon, S. M., Hong, Y.-C., Park, H. J., Lee, J.-E., Kim, S. Y., Kim, J. H., Lee, S.-W., Park, S.-Y., Lee, J. S. and Choi, E. K. (2005), 'The polymorphism and haplotypes of *xrcc1* and survival of non-small-cell lung cancer after radiotherapy', *International Journal of Radiation Oncology Biology Physics* **63**(3), 885–891.
- Yu, Y., Wang, W., Ding, Q., Ye, R., Chen, D., Merkle, D., Schriemer, D., Meek, K. and Lees-Miller, S. P. (2003), 'Dna-pk phosphorylation sites in *xrcc4* are not required for survival after radiation or for v (d) j recombination', *DNA repair* **2**(11), 1239–1252.
- Yuan, P., Liu, L., Wu, C., Zhong, R., Yu, D., Wu, J., Xu, Y., Nie, S., Miao, X., Sun, Y. *et al.* (2010), 'No association between *xrcc1* polymorphisms and survival in non-small-cell lung cancer patients treated with platinum-based chemotherapy', *Cancer biology & therapy* **10**(9), 854–859.
- Yuan, Z., Li, J., Hu, R., Jiao, Y., Han, Y. and Weng, Q. (2015), 'Predictive assessment in pharmacogenetics of *xrcc1* gene on clinical outcomes of advanced lung cancer patients treated with platinum-based chemotherapy', *Scientific reports* **5**, 16482.
- Zaanan, A., Dalban, C., Emile, J.-F., Blons, H., Fléjou, J.-F., Goumard, C., Istanbullu, M., Calmel, C., Alhazmi, K., Validire, P. *et al.* (2014), 'Ercc1, *xrcc1* and *gstp1* single nucleotide polymorphisms and survival of patients with colon cancer receiving oxaliplatin-based adjuvant chemotherapy', *Journal of Cancer* **5**(6), 425.
- Zhang, J., Wu, X.-h. and Gan, Y. (2013), 'Current evidence on the relationship between

- three polymorphisms in the *xrcc7* gene and cancer risk', *Molecular biology reports* **40**(1), 81–86.
- Zhang, J., Zeng, X. T., Lei, J. R., Tang, Y. J. and Yang, J. (2014), 'No association between *xrcc1* gene arg194trp polymorphism and risk of lung cancer: evidence based on an updated cumulative meta-analysis', *Tumor Biology* **35**(6), 5629–5635.
- Zhang, L. and Ma, W. (2014), 'Pharmacogenetics of dna repair gene polymorphisms in non-small-cell lung carcinoma patients on platinum-based chemotherapy', *Genet Mol Res* **13**(1), 228–36.
- Zhang, X. and Miao, X. (2005), 'Polymorphisms in dna base excision repair genes *adprt* and *xrcc1* and risk of lung cancer', *Cancer research* **65**(3), 722–726.
- Zhang, Z. and Wan, J. (2005), 'Genetic polymorphisms in *xrcc1*, *ape1*, *adprt*, *xrcc2*, and *xrcc3* and risk of chronic benzene poisoning in a chinese occupational population', *Cancer Epidemiology and Prevention Biomarkers* **14**(11), 2614–2619.
- Zhao, R. and Chen, G. (2015), 'Role of *gstp1* ile105val and *xrcc1* arg194trp, arg280his and arg399gln gene polymorphisms in the clinical outcome of advanced non-small cell lung cancer', *International journal of clinical and experimental pathology* **8**(11), 14909.
- Zhao, W., Hu, L., Xu, J., Shen, H., Hu, Z., Ma, H., Shu, Y., Shao, Y. and Yin, Y. (2013), 'Polymorphisms in the base excision repair pathway modulate prognosis of platinum-based chemotherapy in advanced non-small cell lung cancer', *Cancer chemotherapy and pharmacology* **71**(5), 1287–1295.
- Zhi, Y., Yu, J., Liu, Y., Wei, Q., Yuan, F., Zhou, X., Song, B., Chen, Z. and Yang, J. (2012), 'Interaction between polymorphisms of dna repair genes significantly modulated bladder cancer risk', *International journal of medical sciences* **9**(6), 498.
- Zhou, C., Ren, S., Zhou, S., Zhang, L., Su, C., Zhang, Z., Deng, Q. and Zhang, J. (2010), 'Predictive effects of *ercc1* and *xrcc3* snp on efficacy of platinum-based chemotherapy in advanced nsccl patients', *Japanese journal of clinical oncology* **40**(10), 954–960.
- Zhou, P.-T., Li, B., Ji, J., Wang, M.-M. and Gao, C.-F. (2015), 'A systematic review and meta-analysis of the association between *ogg1* ser326cys polymorphism and cancers', *Medical Oncology* **32**(2), 31.
- Zhu, D. Q., Zou, Q., Hu, C. H., Su, J. L., Zhou, G. H. and Liu, P. (2015), 'Xrcc1 genetic polymorphism acts a potential biomarker for lung cancer', *Tumor Biology* **36**(5), 3745–3750.
- Zhu, G. and Lippard, S. J. (2009), 'Photoaffinity labeling reveals nuclear proteins that uniquely recognize cisplatin- dna interstrand cross-links', *Biochemistry* **48**(22), 4916–4925.
- Zilfou, J. T. and Lowe, S. W. (2009), 'Tumor suppressive functions of p53', *Cold Spring Harbor perspectives in biology* **1**(5), a001883.

Appendix

A.1 Ethics Committee Report

Institute Ethics Committee Postgraduate Institute of Medical Education and Research, Chandigarh	
Prof. KS Chugh <i>Chairman</i>	Prof. Pallab Ray <i>Convener</i>
	No: PGI/IEC/2014/305 Date: Aug 12, 2014
Chairman Prof. KS Chugh	Dr. Navneet Singh ✓ Department of Pulmonary Medicine, PGIMER, Chandigarh
Members Justice VK Jhanji / Justice MR Agnihotri Dr BN Goswami / Dr Virendra Kumar Dr Girish Varshney Dr Promila Vasudeva Prof. Savita Malhotra Prof. R Kochhar Prof. Savita Kumari Prof. Vanita Suri Prof. S Radhika Prof. Ashim Das Prof. RK Dhiman Dr Jayashree M Dr L Kaman Dr JS Thakur Dr Nusrat Shafiq	Title: Project entitled "Association of single Nucleotide Polymorphisms in the DNA Repair Genes and Survival of Advanced Non-Small-Cell Lung cancer patients treated with platinum-Based Doublet Chemotherapy." Project Number: P-106 Dear Dr Singh, We received five copies of your above mentioned project proposal, which has been assigned the project number P-106 . The project was discussed in the meeting of the Institute Ethics committee held on 09.11.2011. After due consideration and discussion, your project was CONDITIONALLY APPROVED . Subsequently we received the clarification on the issue raised by the Committee and now your project is APPROVED . It is understood that the study will be conducted under your direction strictly as per the submitted protocol. Any deviations from the approved protocol and study documents must be sent to the Ethics Committee for re-approval. Any adverse reaction or condition noted during the study period should be reported to the Ethics Committee immediately. A report of the study should be sent to the Ethics Committee yearly. All future correspondence with respect to this project must mention the project number . Kind Regards Yours sincerely,  (Dr. Pallab Ray) Copy to: Professor-in-Charge, Medical Education Cell, PGIMER, Chandigarh.
Contact Ph: +91 172 2755157 Email: iecpgi@gmail.com	

Office: Department of Medical Microbiology, Second Floor, Research Block A, PGIMER, Chandigarh.

A.2 PATIENT INFORMATION SHEET

PATIENT INFORMATION SHEET

Department of Biotechnology, Thapar University
&
Dept. of Pulmonary Medicine, PGIMER, Chandigarh

Investigator: Dr. Siddharth Sharma, Ph.D

Co-Investigator: Dr. Ashutosh Agarwal, M.D. D.M

Co-Investigator: Dr. Navneet Singh, M.D. D.M

Name of Patient _____

Address (Rural/Urban) _____

C.R. Number _____

Date of Sampling _____

Type of Cancer _____

Stage of Cancer _____

Sex _____

Age (in years) _____

Education _____

Occupation _____

Marital Status _____

Smoking Status

1. Smoker or Non Smoker _____

2. Type of smoking

a) Cigarette b) Beedi c) Hookah d) Cigarette+Beedi e) Beedi+Hookah f) Any other
of smoke

3. No. of Cigarettes/Beedis smoked

a) Per day _____

b) Per week _____

c) Amount of tobacco added to Hookah _____

4. a) Age at which started smoking _____

b) How many years have you smoked _____

5. Time taken to finish cigarette (in min) _____
6. Brand of smoke preferred _____
7. Do you chew tobacco/betelnut
a) Yes b) No
If yes amount chewed _____
8. Any familial case of smoking related cancer in your family: a) Yes b) No
9. Are you only one to smoke in your family

Alcohol Consumption

1. Do you consume alcohol
a) Yes b) No
2. Intake of Alcohol
a) Daily b) 3-5 times a week c) 1-2 times a week d) once a week e) once a month
3. How much alcohol do you consume in ml _____
4. What type of alcohol do you consume _____

PATIENT CONSENT FORM

Name of participant: _____

Name of the Principal (Co-) Investigators: _____

Name of Institution: _____

Name and address of the sponsoring (funding) agency(ies): _____

Documentation of the informed consent

I, _____ have read the information in this form (or it has been read to me). I was free to ask any questions and they have been answered. I am over 18 years of age and, exercising my free power of choice, hereby give my consent to be included as a participant in the project **Association of Single Nucleotide Polymorphisms and Methylation status in the Wnt and AhR signaling pathways with risk for occurrence of Lung Cancer in a North Indian Population**

1. I have read and understood consent form and the information provided to me.
2. I have had the consent document explained to me.

3. I have been explained about the nature of the study.
4. My rights and responsibilities have been explained to me by the investigator.
5. I have the right to not be a subject in the above study and am not willing to give sample.
6. I have informed the investigator of all the treatments I am taking or have taken in the past _____. months including any desi (alternative) treatments.
7. Patient data and confidentiality has been explained to me.
8. I hereby give permission to the investigators to release the information obtained from me as result of participation in this study to the sponsors, regulatory authorities, Government agencies, and ethics committee. I understand that they may inspect my original records.
9. My identity will be kept confidential if my data are publicly presented.
10. I have had my questions answered to my satisfaction.
11. I have decided to be in the research study.

I am aware, that if I have any questions during this study, I should contact the concerned authorities. By signing this consent from, I attest that the information given in this document. I will be given a copy of this consent document.

For adult participants

Name and signature / thumb impression of the participant (or legal representative if participant incompetent): _____ (Name) _____ (Signature)

Date _____ Time _____

Name and signature of impartial witness (required for illiterate patients):

_____ (Name) _____ (Signature)

Date _____ Time _____

Address and contact number of the impartial witness: _____

Name and signature of the Investigator or his representative obtaining consent:

_____ (Name) _____ (Signature)

Date _____

A.3 III

1. 0.5M EDTA: Dissolved 9.306 g of disodium salt of EDTA in 20 ml of deionised water, and then adjusted the pH to 8.0 by 1 M sodium hydroxide. Sterilized the solution by autoclaving.
2. 10% SDS: Dissolved 1 g of SDS in 10 ml of deionised water.
3. 100 mM Tris-Cl (pH 8.0): Dissolved 0.32 g of Tris-Cl in 10 ml of deionised water, then adjusted the pH to 8.0 by 1M sodium hydroxide. Sterilized the solution by autoclaving.
4. 10 mg/ml Proteinase K: Dissolved 10 mg Proteinase K in 1 ml of double distilled water. Sterilized the solution by autoclaving.
5. 1 mg/ml BSA: Dissolved 100 mg of BSA in 100 ml of deionised sterile water and kept at 4 °C overnight.
6. 5 M Sodium chloride (NaCl): Dissolved 5.85 g of sodium chloride in 20ml of deionised water. Sterilized the solution by autoclaving.
7. 5X TBE buffer: Dissolved 54 g of Tris base and 27.5 g of boric acid in 980 ml of double distilled water and then added 20 ml of 0.5 EDTA. Sterilized the solution by autoclaving.
8. Ethidium Bromide (10 mg/ml): Dissolved 1 g of ethidium bromide in 100 ml of water. Mixed the solution properly.
9. Magnesium chloride (MgCl_2) (100 mM): Dissolved 0.41 gms of MgCl_2 in 20 ml of deionised water and sterilized by autoclaving.
10. Sucrose (1 M): Dissolved 3.41 g of sucrose in 10 ml of deionised water and sterilized by autoclaving.
11. TE buffer (pH 8.0): Added 1 ml of 100 mM Tris-Cl (pH 8.0) and 200 μl of 0.5 M EDTA solution to 8.8 ml of deionised water. Sterilized the solution.
12. Triton X- 100 (10%): Took 100 μl of TritonX-100 and mixed with 900 μl of deionised water and mixed properly.