Determinants of Lung Cancer Susceptibility

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences.

2007

Philip Alexander John Crosbie

School of Medicine

ProQuest Number: 11009539

ProQuest 11009539

(EQOFD)

Th29111

Contents

List of Figures

9

List of Tables

Abstract

Lung cancer is the leading cause of cancer mortality globally. Over 90% of cases are caused by chronic exposure to tobacco smoke. However, only one in ten smokers develops the disease. It is postulated that genetic factors modify an individual's susceptibility to tobacco smoke carcinogens and therefore the risk of developing lung cancer. Identification of markers of susceptibility to lung cancer could be used to identify those individuals at most risk. The aim of this project, using a case-control study design, was to investigate determinants of lung cancer susceptibility focusing primarily on DNA repair genotype and phenotype. Subjects were recruited from the Bronchoscopy Unit and Chest Clinic at Wythenshawe Hospital between March 2003 and December 2005. Case status (n=161) was assigned to subjects with an incident diagnosis of lung cancer. Two control groups were recruited, cancer free subjects either from the Bronchoscopy Unit (n=217) or from the Chest Clinic (n=90). All study participants were interviewed by the author prior to diagnosis, using a standardised risk factor questionnaire. Information regarding age, gender, ethnicity, alcohol and cigarette exposure, occupational and medical history was collected. Blood was obtained from all individuals.

Two separate [^{32}P]-labelled oligonucleotide cleavage assays were developed to measure MGMT and MPG DNA repair activities in PBMC extracts. MPG repair activity was found to be significantly greater in cases compared to controls (4.21 ± 1.67 vs. 3.47 ± 1.35 fmole/μgDNA/hr, p=0.005). This was significant in females (p=0.003) but not males (p=0.47). Females with above median MPG activity were six times more likely to have lung cancer than those with below median activity (OR$_{adj}$ 6.33, 95% CI 1.17-34.2). MGMT repair activity was not associated with lung cancer overall, but was significantly higher in current smoking cases compared to controls (p=0.01). The DNA repair activities of MPG and MGMT were significantly correlated (R^2 0.26, p = 0.0001).

All subjects were genotyped for MGMT (K178R + rs12268840), GSTM1, GSTT1 and hOGG1 (Ser326Cys). The MGMT codon 178 polymorphism was not associated with lung cancer; although homozygotes for the MGMT 178R allele had a marked reduction in lung cancer risk which did not reach significance. The MGMT rs12268840 (intron 1) polymorphism showed a significant protective effect against squamous cell carcinoma of the lung (trend test: bronchoscopy = 0.003 and chest = 0.0007 controls). The GSTM1 null genotype was significantly associated with lung cancer when compared to bronchoscopy and chest clinic controls respectively (bronchoscopy OR$_{adj}$ 1.57, 95% CI 1.02-2.43; chest OR$_{adj}$ 1.86, 95% CI 1.09-3.15). A significant dose effect of GSTM1 gene copy number and protection against lung cancer was seen in the chest clinic arm (p=0.02) and to a lesser extent in the bronchoscopy arm (p=0.07). GSTT1 was not associated with lung cancer risk, except when subjects were stratified according to GSTM1 status. hOGG1 genotype was not associated with lung cancer risk.

MGMT allelic expression imbalance was measured in RNA extracted from PBMCs, normal lung and tumour tissue. Imbalance in PBMCs was similar to that seen in lung suggesting a genetic cause. A pilot study measuring DNA oxidation (8-oxodG) and alkylation damage (N7-medG) showed no association with adduct levels and case status, age, gender or smoking in PBMCs. Adduct levels were not correlated but levels of oxidative damage were 20 times greater than alkylation damage, suggesting exposure to oxidising agents may be more important in carcinogenesis than exposure to alkylating agents. Larger studies would be required to further investigate associations between genotypes, enzyme repair activities, specific DNA adduct levels and lung cancer.

Declaration

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of the University of Manchester or any other university or other institute of learning.

Philip A.J.Crosbie

Dedication and Acknowledgements

I would like to give a huge thank you to all my family for putting up with me; especially to Emma, whose advice and endless support kept me going through all the highs and lows of this journey. Much love to Susanna who was delivered by stork during the final year of the project and also to Thomas who was born in the week before the viva; both have given so much joy to a very proud father. I would like to say a big thank you to my Dad who stepped in at the drop of a hat to save the day on too many occasions to count and also to Matthew and Sandra (and bump).

I would like to dedicate this thesis to my Mum.

(1946 – 1998)

Love Always

List of Abbreviations

A	adenine
AAG	alkyladenine DNA glycosylase
AAH	atypical adenomatous hyperplasia
AEI	allelic expression imbalance
APE1	abasic endonuclease 1
APNG	alkylpurine DNA glycosylase
AP site	apurinic / apyrimidinic site
ATP	adenosine 5' triphosphate
BaP	benzo[a]pyrene
BCNU	1,3-bis(2-chloroethyl)-1-nitrosourea
BEC	bronchial epithelial cell
BER	base excision repair
BSA	bovine serum albumin
Bq	becquerel
C	cytosine
CI	confidence interval
CIS	carcinoma-in-situ
COPD	chronic obstructive pulmonary disease
CPD	cigarettes per day
CT	computerised tomography
CV	coefficient of variance
CXR	chest x-ray
CYP	cytochrome P450 enzyme
DNA	deoxyribonucleic acid
DPM	disintegrations per minute
DTT	dithiothreitol
DTE	dithioerytriol
EDTA	ethylenediaminetetra-acetic acid
εA	ethenoadenine
ELISA	enzyme linked immunosorbent assay
ER	oestrogen receptor
FEV1	forced expiratory volume in 1 second
FH	family history
FVC	forced vital capacity
G	guanine
GST	glutathione S-transferase
HPLC-ECD	high performance liquid chromatography-electrochemical detection
Hx	history
MGMT	alkylguanine-DNA alkyltransferase
MMS	methyl methanesulfonate

MMC	mitomycin C
MNU	N-methyl-N-nitroso urea
MPG	N3-methylpurine DNA glycosylase
mRNA	messenger RNA
NER	nucleotide excision repair
N3-meA	N3-methyladenine
N7-meG	N7-methylguanine
N7-medG	N7-methyldeoxyguanosine
NNAL	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN	N-nitrosonornicotine
NSCLC	non-small cell lung cancer
hOGG1	human 8-oxoguanine DNA N-glycosylase
O^6-meG	O^6-methylguanine
OR	odds ratio
8-oxoG	8-oxoguanine
8-oxodG	8-oxodeoxyguanosine
PAH	polycyclic aromatic hydrocarbons
PARP	poly (ADP-ribose) polymerase
PBL	peripheral blood lymphocyte
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PET	positron-emission tomography
PMH	past medical history
PMSF	phenylmethyl sulfonyl fluoride
Pol β	DNA polymerase β
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
SCLC	small cell lung cancer
SD	standard deviation
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SNP	single nucleotide polymorphism
SOC2000	standard occupational classification 2000
SPSS	statistical package for the social sciences
T	thymine
TBE	tris buffered EDTA
TEMED	N,N,N′,N′-tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
TSN	tobacco specific nitrosamine

Chapter One

Introduction

1.1   Lung Cancer Epidemiology

1.1.1   The Impact of Lung Cancer

Lung cancer is the most common cause of cancer related death worldwide, accounting for 1.2 million deaths / year (Parkin, Bray et al. 2005). In the UK, 33 000 people die each year from lung cancer, equivalent to 6% of all cause mortality and one in five of all cancer deaths. Lung cancer is the second most common cancer in males behind prostate cancer and the second most common cancer in females behind breast cancer (Figure 1.1). Patients are rarely diagnosed before the age of 40 years and incidence gradually rises to a peak in the eighth decade (Figure 1.2) (CRUK 2007).

The main risk factor for lung cancer is tobacco smoking. The incidence of lung cancer is directly proportional to smoking prevalence in the population after a lag period of twenty to thirty years (Figure 1.3). In British males, lung cancer incidence rose sharply in the middle and later half of the 20th century reaching a peak in the 1970s. Since this peak rates have been gradually falling with an almost 50% reduction to the present day. This trend is a direct reflection of changes in smoking prevalence in the male population, falling from 70% in the 1950s to 27% in 2002 (UK Government 2004). The incidence of lung cancer in females by comparison gradually increased from the 1970s, reaching a plateau in the early 1990s, since then it has remained fairly constant; a trend that reflects smoking prevalence in the female population. The age standardised incidence rate is almost double in males compared with females (64 vs. 35 / 100 000) but as male incidence falls the difference between the sexes continues to narrow (UK Government 2004).

One striking aspect of lung cancer as a disease is the high mortality and extremely poor survival seen in patients diagnosed. Deaths from lung cancer exceed the combined mortality from breast and colorectal cancer, the next two commonest malignancies (CRUK 2007). In females, deaths from lung cancer exceed those from breast cancer; this is despite the threefold difference in incidence (Figure 1.1). The diagnosis of lung cancer carries one of the worst prognoses of all malignant diseases, with 75% of patients dying within the first year after diagnosis and only 7% five year survival (Coleman, Rachet et al. 2004). One year survival rates have improved from 13-15% in 1971-5 to 25-26% in 2000-1, unfortunately no significant change in long term survival has occurred over this time (Coleman, Rachet et al. 2004).

a) Males
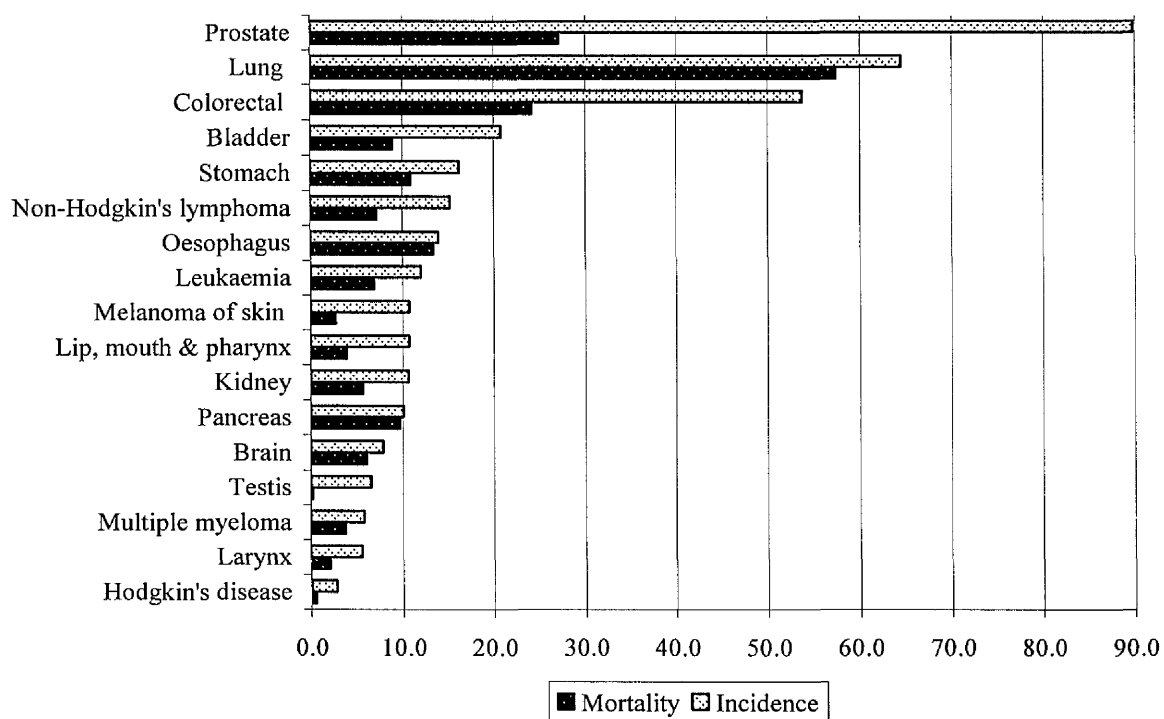
b) Females
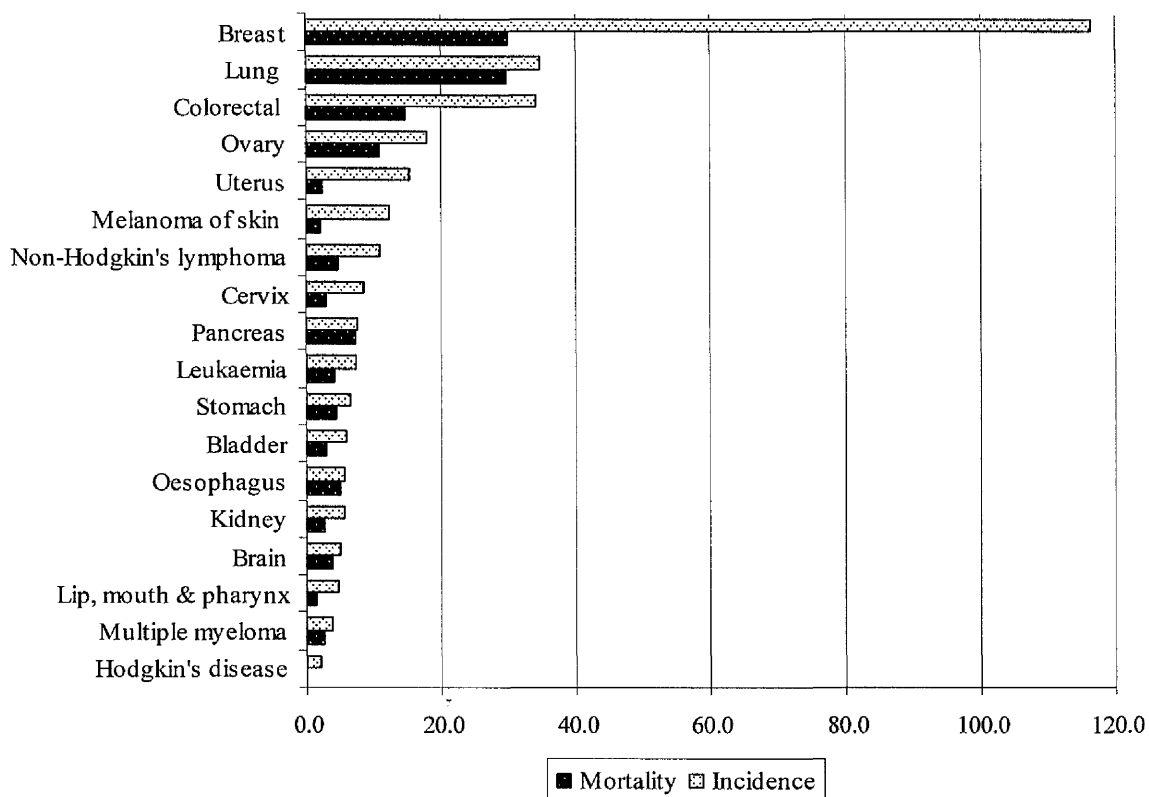
Figure 1.1 Major UK cancers (three year average 2001-2003): age standardised incidence and mortality rates / 100 000 population by gender (UK Government 2006).
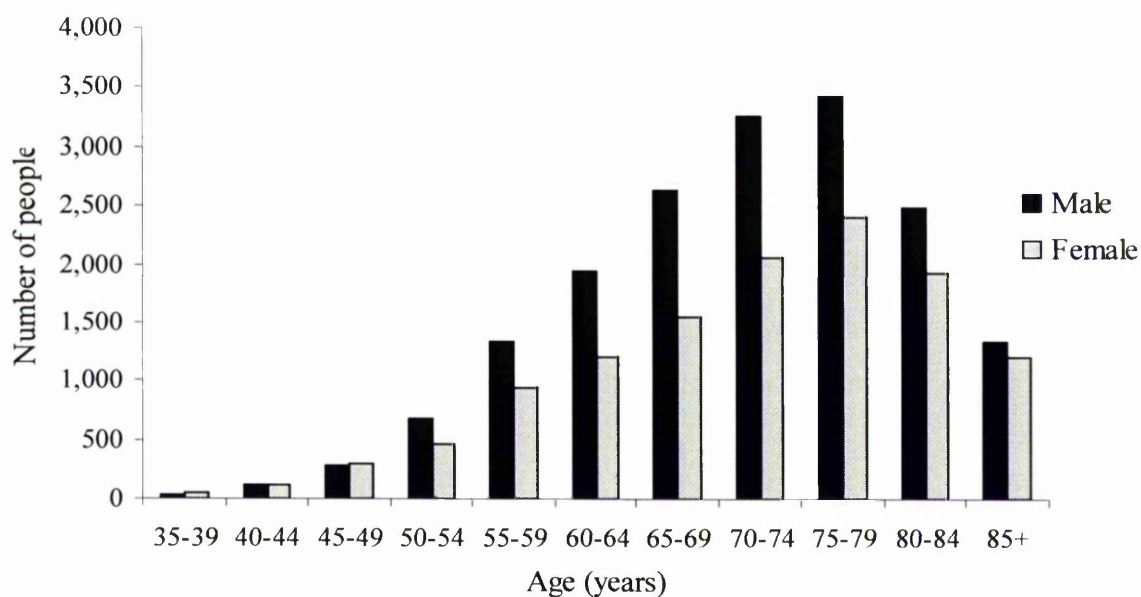
Figure 1.2 Age specific incidence of lung cancer according to gender (England, 2003). (CRUK 2007).
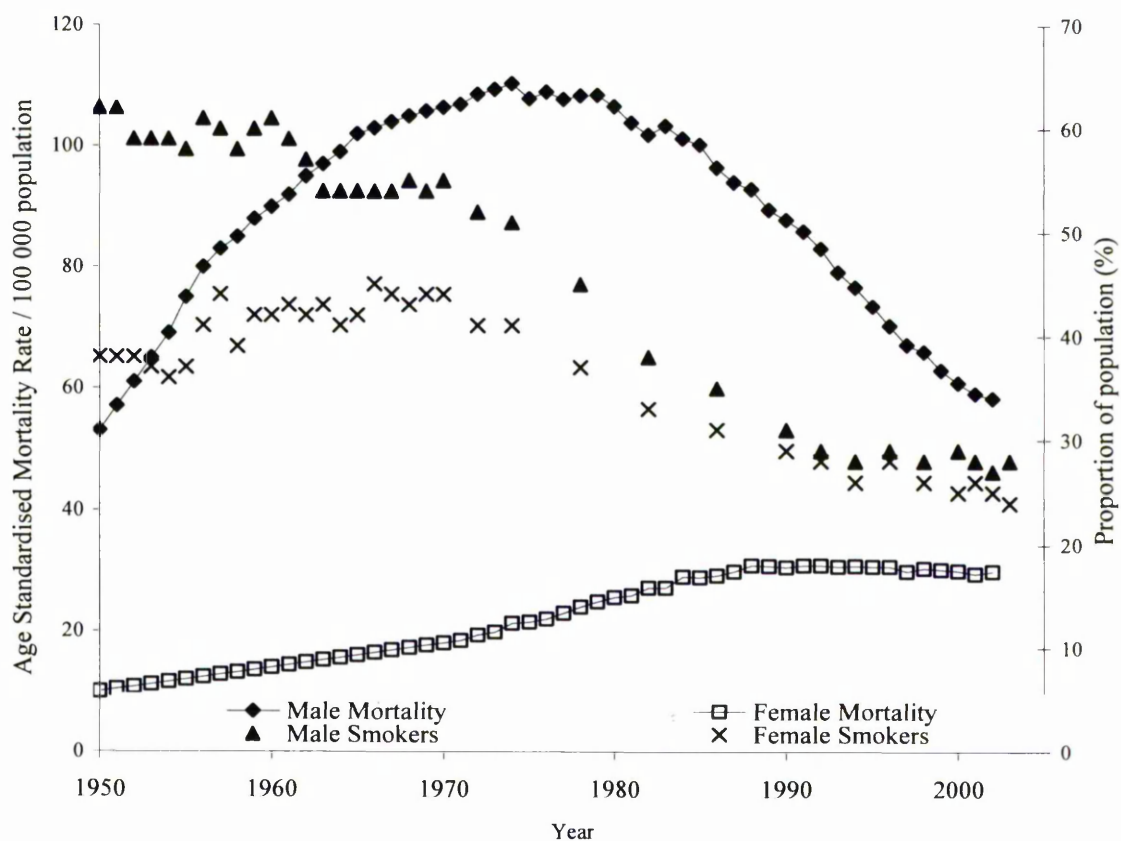
Figure 1.3 Age standardised mortality rate and smoking prevalence in the UK according to gender (1950-2003). (CRUK 2007).

1.1.2 Risk Factors for Lung Cancer

1.1.2.1 Tobacco Smoke Exposure

Epidemiological evidence of an association between smoking and lung cancer was first demonstrated in the 1950s (Doll and Hill 1950; Doll and Hill 1952). Smoking was shown to increase the risk of developing lung cancer by at least ten fold compared to a lifelong non-smoker. The risk of lung cancer increases in proportion to both the number of cigarettes smoked and the number of years a person has smoked. However, the duration of cigarette smoking is thought to be a more important factor (Flanders, Lally et al. 2003). It is estimated that tobacco smoking is responsible for 90% of cases (Alberg and Samet 2003); in one study from Edinburgh, of all the patients diagnosed with lung cancer only 2% were lifelong non-smokers (Capewell, Sankaran et al. 1991).

The identification of chronic tobacco smoke exposure as the main modifiable risk factor for lung cancer, places this disease in a unique position amongst the major cancers, as being eminently preventable. Despite this knowledge global tobacco production and consumption continue to rise, although a disparity exists between the developing and developed world. Increasing demand for tobacco products is seen in the developing world fuelled by increasing populations and increased wealth. The consumption of cigarettes in China for example has quadrupled over the past three decades. In contrast, the consumption of tobacco products is gradually decreasing in the developed world due to a combination of factors such as increased taxation on cigarettes and increased public awareness of the detrimental effects of smoking (Nations 2003). In the UK, the overall prevalence of smoking has decreased over time (Figure 1.3). The rate of reduction in smoking was at its greatest in the 1970s, this rate slowed during the 1980s and has remained fairly constant over the last decade. In 2002, 27% of men and 25% of women were classified as current smokers (UK Government 2004). The prevalence of smoking is age dependent, reaching a peak in the 20 to 24 age category, where 40% of men and 35% of women smoke (Figure 1.4). Smoking then decreases with increasing age to a prevalence of 16% in the over 60 age group. The number of women smoking has changed to a lesser extent than in men, in the under 20 age range for instance more females smoke than males.

25
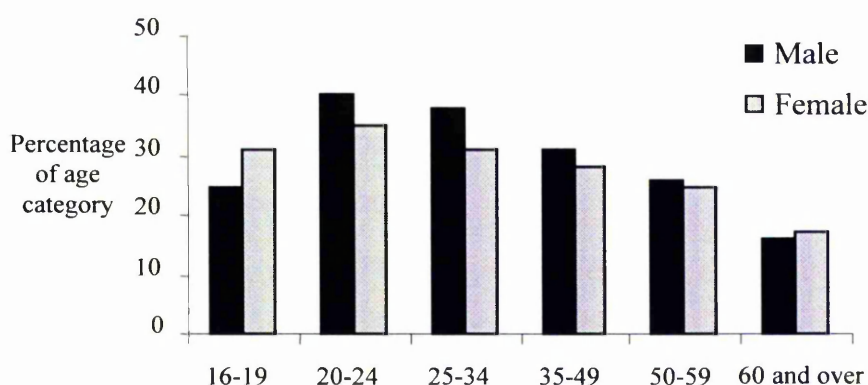
Figure 1.4 Age specific prevalence of smoking. (UK, 2001).

In 1990, the cumulative risk of dying from lung cancer in the UK, for current smokers up to the age of 75, was 15.9% in men and 9.5% in women (Peto, Darby et al. 2000). An observational population-based cohort study in 19 714 Danish smokers, showed a 27% reduction in lung cancer in smokers (>15 cpd) who were able to reduce tobacco smoke exposure by 50% (Godtfredsen, Prescott et al. 2005). The risk can be reduced further by smoking cessation; the reduction is proportional to the length of time since stopping smoking but never drops to non-smoking levels (Figure 1.5). Stopping before middle age avoids more than 90% of the risk attributable to tobacco and reduces the incidence of all histological types of lung cancer with the greatest reduction seen in small and squamous cell carcinomas (Tong, Spitz et al. 1996; Khuder and Mutgi 2001). There are twice as many former smokers than current smokers over the age of fifty in the UK, lung cancers are therefore commonly diagnosed in patients who no longer smoke.
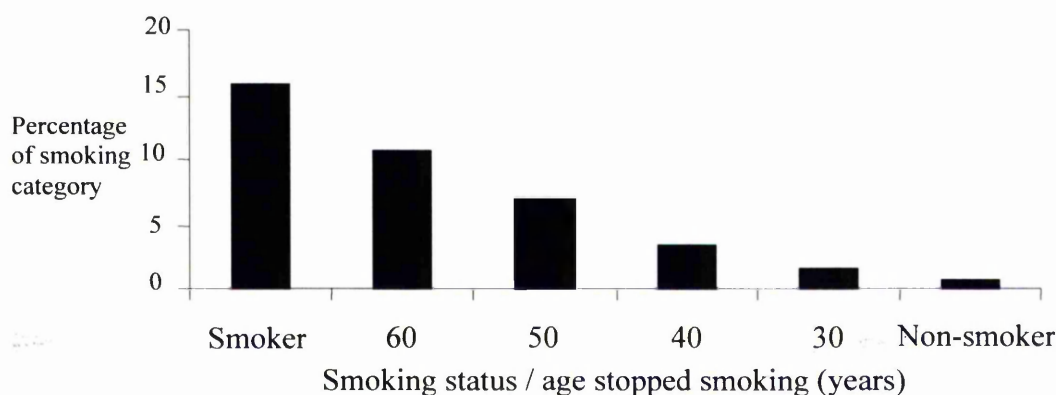
Figure 1.5 The cumulative risk of lung cancer mortality in British male: smokers, ex-smokers and lifelong non-smokers up to age 75.

Exposure to tobacco smoke is most commonly through the consumption of cigarettes. The use of other tobacco products has also been linked to an increased risk of lung cancer. Cigar smoking increases, by a factor of between two to five fold, the relative risk of developing lung cancer, when compared to non-smokers (Iribarren, Tekawa et al. 1999; Shapiro, Jacobs et al. 2000). Current pipe smoking also increases the risk of lung cancer up to five fold compared to never smokers (Henley, Thun et al. 2004). Stopping smoking and switching to smoking cigars or a pipe halves the combined risk of dying from lung cancer, ischaemic heart disease and chronic obstructive pulmonary disease (Wald and Watt 1997). However the risk remains 50% higher than a lifelong non-smoker.

Exposure to environmental tobacco smoke, also known as passive smoking, is the involuntary exposure of non-smokers to tobacco from the smoking of others. The magnitude of exposure is much less than that for active smoking, but environmental tobacco exposure can occur earlier in life and over a long duration. A meta-analysis of 37 published studies, involving 4626 cases of lung cancer in non-smokers, estimated an excess in risk of developing lung cancer of 24% (CI 13-36%) in a non-smoking spouse living with a smoker (Hackshaw, Law et al. 1997). There was no identified bias or confounding factor that could explain the increase in risk and a dose response was observed. However, the exact risk attributable to environmental tobacco smoke remains controversial (Copas and Shi 2000).

1.1.2.2 Occupational and Environmental Exposures

It is estimated that 15% of lung cancer in males and 5% in females could be attributed to occupational exposures (Coultas and Samet 1992). Agents encountered in industry accepted as pulmonary carcinogens include arsenic, chloromethyl ethers, chromium, nickel, polyaromatic hydrocarbons and vinyl chloride. Lung cancer has been associated with the coal-gas industry, metal refining and smelting (Brambilla and Corrin 2003). Exposure to fossil fuel combustion products can include carcinogens such as polycyclic aromatic hydrocarbons and metals e.g. arsenic, nickel and chromium. A higher risk of lung cancer has been reported in urban when compared to rural areas, after allowing for the effects of smoking (Dockery, Pope et al. 1993) suggesting that air pollution might also have an effect on lung cancer incidence.

Asbestos is a well known occupational carcinogen, exposure to which can lead to lung cancer and mesothelioma. The risk of lung cancer associated with asbestos exposure is dose dependent and varies according to the type of asbestos fibre. The risk is considerably higher after exposure to amphibole fibres (crocidolite and amosite) than chrysotile fibres. Heavy asbestos exposure on its own leads to a five fold increase in lung cancer mortality when compared to normal controls. When asbestos and smoking exposure are present together, they combine synergistically to increase the risk of lung cancer mortality considerably (up to 53 fold over normal controls (Alberg and Samet 2003)). There is some controversy as to whether asbestos is directly carcinogenic or acts indirectly by causing asbestosis which eventually leads to cancer (Weiss 1999). It is estimated that asbestos accounts for 2-3% of lung cancer in the UK (Darnton, McElvenny et al. 2006).

Radon is a decay product of uranium-228 and radium-226 and is capable of damaging the respiratory epithelium via the emission of alpha particles. The carcinogenic effect of radon was first noted in studies on underground miners (Grosche, Kreuzer et al. 2006). Radon is known to accumulate in houses but at concentrations that are much lower than those found in mines. A meta-analysis of eight studies, involving 4263 cases of lung cancer, concluded that the relative risk of being diagnosed with lung cancer was 1.14 times greater (CI 1.0-1.3) for a long term resident in a house with a radon concentration of 150 Bq/m^3. This concentration is equalled or exceeded in 6% of American homes (Lubin and Boice 1997). One review estimated that residential radon exposure was responsible for 2000 lung cancer deaths per year in the UK (Darby, Hill et al. 2001).

1.1.2.3 Pre-existing Medical Conditions

The presence of benign respiratory disease can increase the incidence of lung cancer in affected patients. The incidence of lung cancer is increased in patients with cryptogenic fibrosing alveolitis (Turner-Warwick, Lebowitz et al. 1980); the relative risk of lung cancer in one study was 7.3 when compared to normal controls, this effect was independent of smoking (Hubbard, Venn et al. 2000). The risk of lung cancer has also been shown to be significantly increased in patients with obstructed lung disease, including emphysema (Kishi, Gurney et al. 2002) and in non-smoking asthmatics (Santillan, Camargo et al. 2003).

The incidence of a second primary malignancy is significantly increased in patients with a history of head and neck cancer. In one study, 28% of patients developed a second primary neoplasm (Holland, Arsanjani et al. 2002). More patients died as a result of a second malignancy than their primary disease. The reported prevalence of lung cancer diagnosed during follow up was 6.9-11%, with an annual incidence of 1-2% (Deleyiannis and Thomas 1997; Holland, Arsanjani et al. 2002).

1.1.2.4 Dietary Factors

A high dietary intake of fruit and vegetables has been associated with a reduction in the risk of developing lung cancer (Feskanich, Ziegler et al. 2000; Smith-Warner, Spiegelman et al. 2003; Miller, Altenburg et al. 2004). Elevated serum levels of α-tocopherol the major derivative of vitamin E has also been shown to be associated with a reduction in lung cancer risk (Woodson, Tangrea et al. 1999). However, several large placebo controlled double blind trials investigating the effect of dietary supplementation with anti-oxidants (vitamin E and β-carotene) showed no protective effect against the development of lung cancer (The Alpha-Tocopherol 1994; Hennekens, Buring et al. 1996; Omenn, Goodman et al. 1996). Indeed, there was an *increase* of 18% in lung cancer diagnoses in patients given β-carotene in one Finnish study (The Alpha-Tocopherol 1994).

The consumption of cruciferous vegetables, which are rich in isothiocyanates, has a protective effect against lung cancer in subjects null for both GSTT1 and GSTM1 (Brennan, Hsu et al. 2005). Isothiocyanates have been shown in animal models to be strong inhibitors of lung cancer (Hecht 1996). Subjects positive for both genes were not protected possibly because GSTM1 and GSTT1 encode enzymes responsible for the elimination of isothiocyanates (London, Yuan et al. 2000).

1.1.2.5 Genetic Factors

The evidence for an environmental aetiology, namely tobacco, for lung cancer is unequivocal; a person is highly unlikely to be diagnosed with lung cancer if they are a lifelong non-smoker without a significant occupational history (Thun, Henley et al. 2006). However, there is consistent evidence that lung cancer clusters within families. For example, a study by Tokuhata and Lilienfeld (1963) showed a threefold increase in lung

cancer incidence among parents and siblings of lung cancer cases than among relatives of the patient's spouse (Tokuhata and Lilienfeld 1963). A recent meta-analysis of 31 case-control, 17 cohort and 8 twin studies showed a two fold increased risk of lung cancer for subjects with a positive family history, the risk was greatest when lung cancer was diagnosed at a young age and with the presence of multiple affected family members (Matakidou, Eisen et al. 2005). However, the authors were cautious about their findings because of the difficulty in separating common environmental exposures from genetic factors in family aggregates.

The molecular basis for familial increased risk is unclear. Ooi *et al*, showed an increased risk of lung cancer in relatives of cases when compared to controls (odds ratio 3.2) after correcting for the confounding effects of age, gender, smoking history and occupational risk factors (Ooi, Elston et al. 1986). Analysis of this study, investigating the possible mode of inheritance of genetic susceptibility to lung cancer, suggested that Mendelian codominant inheritance was the most likely (Sellers, Bailey-Wilson et al. 1990). This putative gene was estimated to be responsible for 69% of lung cancer at age 50, 47% at 60 and 22% at age 70. The peak incidence of lung cancer occurs in the 70 to 80 age range; therefore total numbers attributable to this gene are low. A more recent study of families with multiple affected members identified a major susceptibility locus on chromosome 6q that was acting in an autosomal dominant or co-dominant manner, modifying lung cancer risk (Bailey-Wilson, Amos et al. 2004).

The hypothesis that a genetic predisposition to lung cancer is inherited, but the trait is only expressed in the presence of a major environmental exposure, such as tobacco smoke, is a distinct possibility. It is likely that the risk of lung cancer for individuals is determined through the interaction of multiple genes each having a small effect and environmental exposure.

1.2   Clinical Features of Lung Cancer

1.2.1   Presenting Symptoms and Diagnosis

The clinical presentation of lung cancer is extremely varied (Table 1.1) (Beckles, Spiro et al. 2003). Symptoms can be caused locally by the primary tumour, secondary to metastatic spread or be non-specific such as weight loss. These are common to a multitude of other respiratory and cardiac diseases and none is a specific indicator for lung cancer; this can lead to diagnostic delay. In one Scottish study, the median time between the onset of the first symptoms and diagnosis was 93 days, this ranged between 0 to 1118 days (Group 1987).

Table 1.1 Initial symptoms and signs of lung cancer (Beckles, Spiro et al. 2003).

Symptoms and signs	Range of frequencies (%)
Cough	8-75
Weight loss	0-68
Dyspnoea	3-60
Chest pain	20-49
Haemoptysis	6-35
Bone pain	6-25
Clubbing	0-20
Fever	0-20
Weakness	0-10
Superior vena cava obstruction	0-4
Dysphagia	0-2
Wheezing and stridor	0-2

A study by Carbone *et al*, showed that although 27% of patients presented with symptoms relating to the primary tumour, the majority (59%) presented with either non-specific symptoms such as weight loss and fatigue or specific symptoms relating to metastatic disease (Carbone 1970). Prognosis was related to the source of symptoms at diagnosis. The five year survival for asymptomatic patients was 18% in the same study, falling to 12% for patients presenting with symptoms relating to the primary tumour, 6% for those with non-specific symptoms and 0% for symptoms specific to metastatic disease.

The patient's presenting symptoms, physical examination and chest X-ray (CXR) are used to formulate a differential diagnosis. An abnormal CXR can be extremely valuable in directing the diagnostic process and also providing information about the extent of the disease (Figure 1.6). The chances of having lung cancer are much lower in patients with a normal CXR. In one report of 900 lung cancers only 15 (1.7%) had a normal CXR at presentation (Fletcher, Johnston et al. 1976). The earliest a tumour can be detected on a CXR is when its diameter reaches 1cm. At this stage the tumour may have been present for between 2.4 and 13.2 years, depending on the histological subtype (Table 1.2) (Geddes 1979).

Table 1.2 The natural history of untreated lung cancer (Geddes 1979).

Histological Type of Lung Cancer	Volume doubling-time (days)	Years from malignant change to:		
		Earliest diagnosis (1cm)	Usual Diagnosis (3cm)	Death (10cm)
Squamous	88	7.1	8.4	9.6
Adenocarcinoma	161	13.2	15.4	17.6
Undifferentiated	86	7.1	8.2	9.4
Small-cell	29	2.4	2.8	3.2

Radiographic abnormalities can only lead to a presumptive diagnosis, for a full assessment of the patient it is necessary to obtain histological or cytological samples and stage the disease using computed tomography (CT) and positron emission tomography (PET). Obtaining tissue for diagnostic purposes can be achieved by non-invasive methods such as sputum cytology or more invasive methods including bronchoscopy (Figure 1.7). Flexible bronchoscopy is an extremely useful tool in the diagnosis of lung cancer. It is most effective for central tumours, where direct forceps biopsy combined with washing and brushings is a very sensitive technique for obtaining tissue samples (88% sensitivity). The sensitivity falls for the diagnosis of peripheral lesions to 69% (Schreiber and McCrory 2003). Bronchoscopy is valuable not only for the procurement of tissue samples, but also in the assessment of disease extent in the central airways. The bronchoscope can also be used for early diagnosis (autofluorescence) and treatments such as brachytherapy, photodynamic therapy and airway stenting.
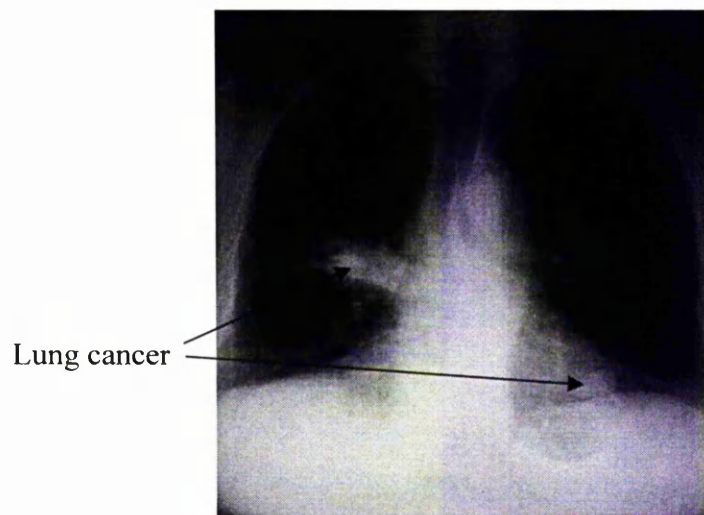
Figure 1.6 The chest x-ray appearance of a primary lung carcinoma.

a) The author holding a bronchoscope

b) The bronchoscopic appearance of lung cancer (in the right main bronchus)
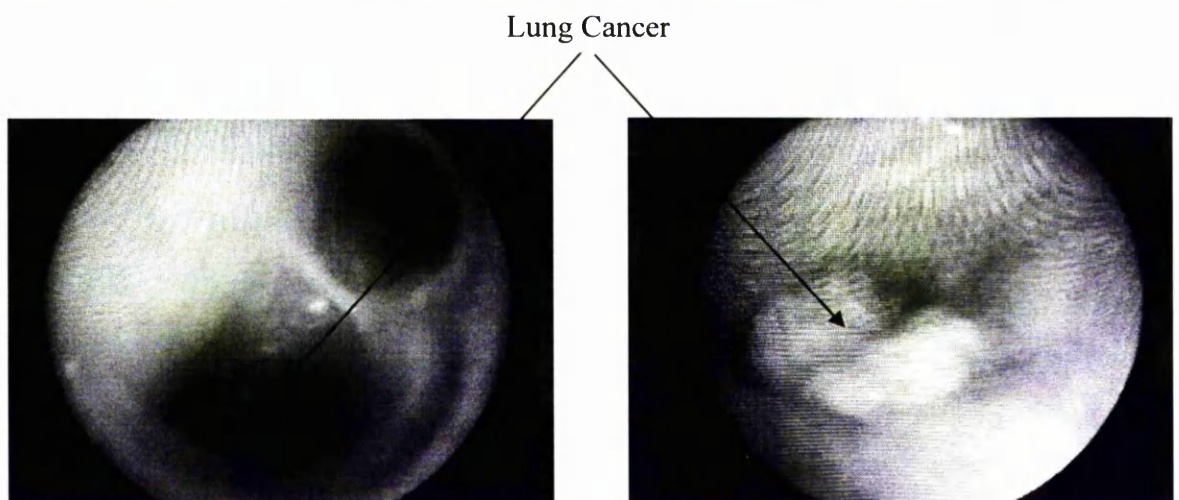
Figure 1.7 The (a) bronchoscope and (b) bronchoscopic appearance of lung cancer.

1.2.2 Histological Types

Lung cancer originates from the epithelial lining of the lung and is classified according to four main cell types that account for 95% of the total incidence. These are squamous cell carcinoma, adenocarcinoma, large cell carcinoma and small cell carcinoma.

Squamous cell carcinoma arises centrally in the lung in 60 to 80% of cases. Histologically they show keratinisation and / or intercellular junctions (prickles). The development of squamous cell carcinoma is thought to progress through several stages from basal cell hyperplasia, metaplasia, dysplasia and carcinoma-in-situ (Greenberg, Yee et al. 2002). It is not known what proportion of lesions progress at each stage and all may be reversible (Chyczewski, Niklinski et al. 2001). Progression from carcinoma-in-situ to carcinoma occurred in 56% of lesions in one study (Venmans, van Boxem et al. 2000).

Most adenocarcinomas occur in the periphery of the lung (75%). They show either gland or intracytoplasmic mucin formation. Adenocarcinoma is the most common cell type found in lung tumours of lifelong non-smokers (42%) (Capewell, Sankaran et al. 1991). The precursor to peripheral parenchymal adenocarcinoma is thought to be atypical adenomatous hyperplasia (AAH) (Greenberg, Yee et al. 2002). The natural history of this lesion in a patient without lung cancer is not known. Bronchoalveolar carcinoma is a subtype of adenocarcinoma that shows a lepidic growth pattern (growth along intact alveolar septa). This form of lung cancer has the lowest association with smoking.

Large cell carcinoma is defined by the exclusion of glandular or squamous differentiation under light microscopy. It includes all poorly differentiated non-small cell carcinomas that are not further classifiable under routine examination. They are usually peripheral masses at presentation and characterised histologically by sheets of round to polygonal cells with prominent nucleoli and abundant cytoplasm.

Small cell tumours account for approximately 20% of lung cancers. They are composed of a pleomorphic population of small cells that are fragile and often necrotic. Small cell is part of a larger category of lung tumours with neuroendocrine properties. These tumours have the ability to synthesise neuropeptides that can manifest as paraneoplastic syndromes. The majority (60%) present as a perihilar mass. It has an aggressive clinical course with patients often presenting with widespread metastatic disease.

There is no significant difference in therapeutic approach between squamous cell, large cell and adenocarcinomas; these types are therefore grouped together under the term Non Small Cell Lung Carcinoma (NSCLC). Small Cell Lung Carcinoma (SCLC) is a more aggressive lesion and is sensitive to chemotherapy, it is therefore more important for treatment purposes to differentiate between small cell and non-small cell than identifying the individual subtype.

a) English males (1978–1999)
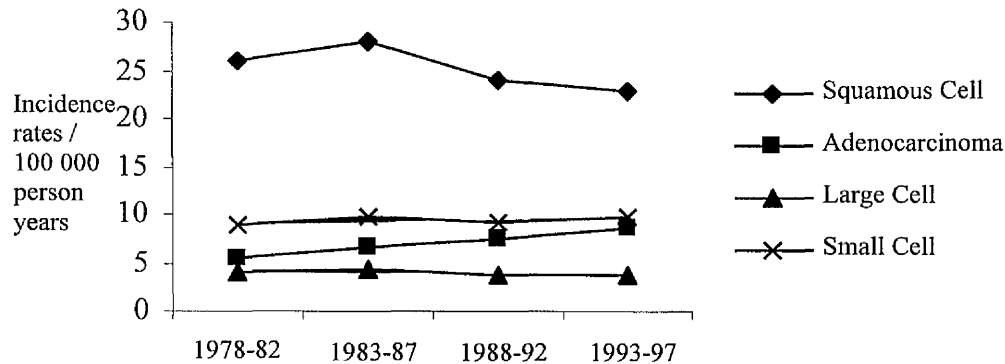
b) English females (1978-1999)
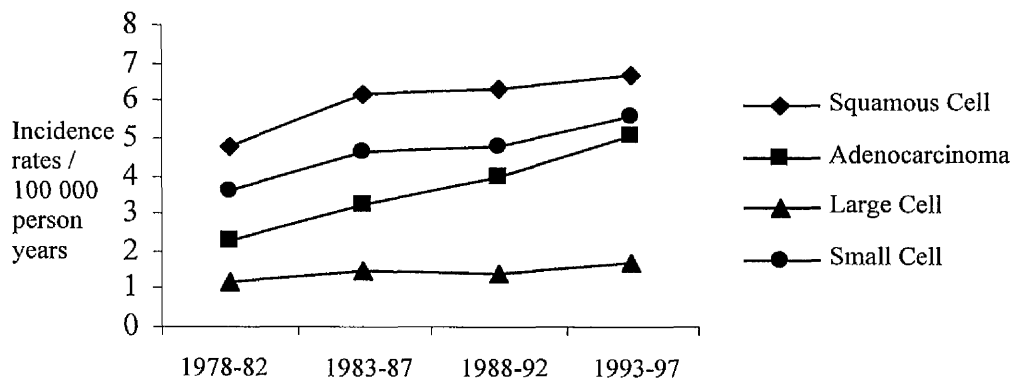
Figure 1.8 Age standardised incidence rates of lung cancer histology / 100 000 person years in English: a) males and b) females (Janssen-Heijnen and Coebergh 2003).

The frequency of each subtype varies between geographical area and according to gender. A report from the Edinburgh Lung Cancer Group looking at the frequencies of different histological types of lung cancer, from 1981 to 1984, showed that squamous cell carcinoma was the most common type in both sexes accounting for 52% in males and 38% in females (Group 1987). Two studies in the USA have shown a marked change in the incidence of histological types of lung cancer over the past three decades (Vincent, Pickren

et al. 1977; el-Torky, el-Zeky et al. 1990). Adenocarcinoma has increased dramatically over this time to become the most common cell type in women and second most common in men. The incidence of adenocarcinoma has also been increasing across Europe, including England in both males and females (Figure 1.8). In 1978 the proportion of adenocarcinomas diagnosed in males was 12% of all lung cancers rising to 19% in 1998; this trend was also seen in females where the proportion rose from 19% in 1978 to 27% in 1997 (Janssen-Heijnen and Coebergh 2003).

1.2.3 Staging

The investigation and staging of patients with lung cancer is directed towards treating the individual with the most appropriate therapy. The TNM staging system (Tumour, Nodes, Metastases) is used in NSCLC to identify patients suitable for surgery; as this is the most effective treatment. Stages 1, 2 and 3A are deemed operable but with distinct survival rates; 3B and 4 are inoperable stages (Mountain 1997). The TNM staging protocol is not applicable to patients diagnosed with small cell lung carcinoma; the high incidence of lymph node and distant metastatic spread at presentation signifies that the TNM classification has no bearing on prognosis. The staging of small cell disease has been simplified into two groups. Limited disease is confined to one hemithorax and the ipsilateral supraclavicular fossa. Extensive disease is any disease outside this area.

1.2.4 Treatment

There are multiple factors that determine prognosis in lung cancer, the most important of which is the stage of disease at diagnosis. The specific treatment of a patient will depend on factors such as cell type, stage and the patient's general medical condition / performance status. In NSCLC, early stage disease can be effectively treated with surgical resection and five year survival in excess of 60% has been reported (Mountain 1997). Indeed survival after surgical resection of small stage 1A tumours has been reported to be as high as 85.5% (Gajra, Newman et al. 2003). Patients with early stage disease who refuse or are not medically fit enough for surgery have a survival benefit if treated with radiotherapy (Sirzen, Kjellen et al. 2003).

Unfortunately the majority of patients present with advanced disease, stages 3B and 4, at which point treatments are much less effective. This is reflected in survival rates of 1 to 5% in these stages (Mountain 1997). Surgery or radical radiotherapy for the majority of patients is therefore not an option. Palliative chemotherapy has been shown to improve survival in advanced NSCLC when compared to best supportive care (1995).

The more aggressive nature of small cell carcinoma often results in patients presenting with widely disseminated disease. Untreated, patients rarely survive more than a few months. Chemotherapy, dramatically improves survival when compared to best supportive care (Agra, Pelayo et al. 2003). Prognosis is related to extent of disease at diagnosis; response to chemotherapy in extensive disease is often brief and median time to death is 9 months (Chute, Chen et al. 1999). Long term survival of 12% in limited stage disease can be achieved with chemoradiotherapy (Janne, Freidlin et al. 2002). Multiple chemotherapeutic agents are in use but a combination therapy with a platinum based compound e.g. cisplatin is currently recommended (Pujol, Carestia et al. 2000).

1.2.5 Lung Cancer Screening

Lung cancer has a long pre-clinical phase during which time the tumour may grow without causing symptoms (Table 1.2). Diagnosis prior to the development of symptoms might result in earlier diagnosis and better outcome. The diagnosis of asymptomatic patients requires the use of screening techniques. The effectiveness of lung cancer screening has been studied in males during the 1970s. Screening for lung cancer using serial chest X-rays (CXRs) (± sputum cytology), was shown not to reduce disease specific mortality (Melamed, Flehinger et al. 1984; Fontana, Sanderson et al. 1986; Kubik and Polak 1986; Tockman 1986). The Mayo Lung Project (Fontana, Sanderson et al. 1986) is to date the most influential trial to have investigated lung cancer screening. Prevalence screening, using CXR and sputum cytology diagnosed 91 cases of lung cancer from 10933 male study participants (0.83%). Subjects were randomised to a screened group (n = 4618) and a control group (n = 4593). The screen consisted of a four monthly CXR and sputum sample; the control group received one off advice to undergo an annual CXR and sputum cytology. Screening lasted for six years prior to three years follow up. The number of lung cancers detected was 206 in the study group and 160 in the control group. There were 43 more cases of resectable lung cancer in the study group but also 3 more cases of unresectable

cancer. The overall 5 year lung cancer survival in the screened group was 33% which compared with 15% in the control. In the screened group there were 122 lung cancer deaths and 115 deaths in the control. The lung cancer mortality rate was 3.2 and 3.0 / 1000 person years respectively.

All of these trials showed advantages to screening in stage distribution, resectability, survival and fatality (Fontana, Sanderson et al. 1991). The trials have received criticisms concerning study design, statistical analysis, contamination, inherent biases and the low sensitivity of CXRs (Melamed, Flehinger et al. 1984; Patz, Goodman et al. 2000; Ellis and Gleeson 2001). However, no mortality improvement was shown in either the Mayo or Czech trials. Mortality was used as the 'gold standard' to avoid potential biases that are known to affect other measures of screening outcomes (lead time, length and overdiagnosis bias) (Parkin and Moss 2000). However, mortality is directly related to the incidence of cancer diagnosed which was increased in all the screened groups. If the increased incidence of lung cancer in these groups was due to overdiagnosis bias then all that screening achieved was the early discovery of clinically irrelevant tumours.

The arguments against the discovery of indolent lung tumours or overdiagnosis bias come from several sources. Firstly, from the screening studies themselves; approximately 70% of stage I patients who were treated surgically survived more than five years, compared to only two (out of 45) five-year survivors who were not operated on (Flehinger, Kimmel et al. 1992). Similar evidence comes from a retrospective study by McGarry *et al*, who looked at survival in 128 patients with stage 1 and 2A lung cancer. The patients received either surgery, radiotherapy or no treatment. The median survival for patients receiving no treatment was 14.2 months, for patients treated with radiotherapy 19.9 months (there was no difference between curative and palliative intent) and for patients who were treated surgically 46.2 months. The untreated group had a very poor outcome with > 50% dying of lung cancer (McGarry, Song et al. 2002). Secondly, the extent of surgery has a significant impact in patient outcome for stage I disease. (Dominioni, Imperatori et al. 2000; Sugarbaker and Strauss 2000). The results of an American prospective randomised trial, comparing lobectomy to limited resection, showed a 75% increased recurrence rate, 30% increased overall death rate and 50% increased death from cancer rate in the limited resection group (Ginsberg and Rubinstein 1995). This further underlines the malignant potential of even early stage disease and the importance of early detection.

The major tool used in the first screening programmes was the CXR. The sensitivity of CXRs in detecting small tumours is estimated to be 16% (Flehinger, Kimmel et al. 1992). In one retrospective study 18 radiologists failed to detect 27 potentially resectable bronchogenic carcinomas (mean diameter 1.6cm +/- 0.8) (Austin, Romney et al. 1992). In the Mayo project 90% of 50 peripheral lung cancers were identified on CXRs taken four months previously (Muhm, Miller et al. 1983). The availability of CT scanning to increase the sensitivity of lung cancer detection may well have a major impact on lung cancer screening. The effectiveness of CT compared to CXR can be seen in the baseline results of the Early Lung Cancer Action Project (ELCAP) (Henschke, McCauley et al. 1999). This study enrolled 1000 asymptomatic current or former smokers over the age of 60. The prevalence of lung cancer detected by CT was 2.7% compared with 0.7% by CXR. Of 27 detected lung cancers, 26 were resectable. Sone *et al*, showed that CT picked up 10 times as many lung cancers than CXR in a general population screen (Sone, Takashima et al. 1998).

The prevalence of lung cancer is dependent on the population studied. Austin *et al* (2002) calculated the prevalence of stage 1 lung cancer in patients undergoing assessment for surgical management of emphysema at 3% (stages II-IV 0.7%) (Austin, Pearson et al. 2002) which compares to 0.48% reported by Sone *et al* (Sone, Takashima et al. 1998). Lung cancer most commonly manifests as a non-calcified pulmonary nodule. In the ELCAP study nodules were detected in 23% of participants, of these approximately 10% had a malignant lesion. The monitoring of growth in nodules that were not readily identified as benign or malignant initially resulted in 28 biopsies being taken; only one of which was for a benign lesion. However, the management of these nodules is a very significant problem in terms of diagnosis of lung cancer and the increased number of scans required to follow these lesions up. Increased access and use of positron emission tomography which if used in combination with CT scanning has recently been reported to have a negative predictive value of up to 100%, will help in the management of detected nodules (Bastarrika, Garcia-Velloso et al. 2005). The results of two ongoing studies of lung cancer screening using both CXRs and high resolution CT scanning will add to the debate over the next few years (Truong and Munden 2003; Oken, Marcus et al. 2005).

1.3 Tobacco Smoke and Lung Cancer

Chronic exposure to tobacco smoke is known to be the major risk factor for the development of lung cancer (Doll and Hill 1950; Doll, Peto et al. 2004). The link between smoking and lung cancer is thought to be the chronic exposure of target cells within the lung to carcinogens present in tobacco smoke (Figure 1.9). These carcinogens, once activated are then able to induce genetic and epigenetic changes in DNA, an early and essential step in the initiation of carcinogenesis (Hecht 2002). This damage is often in the form of DNA adducts; these are covalent addition products formed between reactive chemical species and sites within the DNA molecule. One consequence of DNA damage in key groups of genes known as tumour suppressor genes or oncogenes includes the loss of the normal regulatory controls governing cellular proliferation which is a central feature of cancer cells (Hanahan and Weinberg 2000).

CARCINOGEN EXPOSURE

⬇

CARCINOGEN ABSORPTION

⬇

CARCINOGEN ACTIVATION   ⟶   DETOXIFICATION / EXCRETION

⬇

DNA DAMAGE (e.g. DNA Adducts)   ⟶   DNA REPAIR / APOPTOSIS

⬇

MUTATIONS in KEY GENES (e.g. Tumour Suppressor Genes and Oncogenes)

⬇
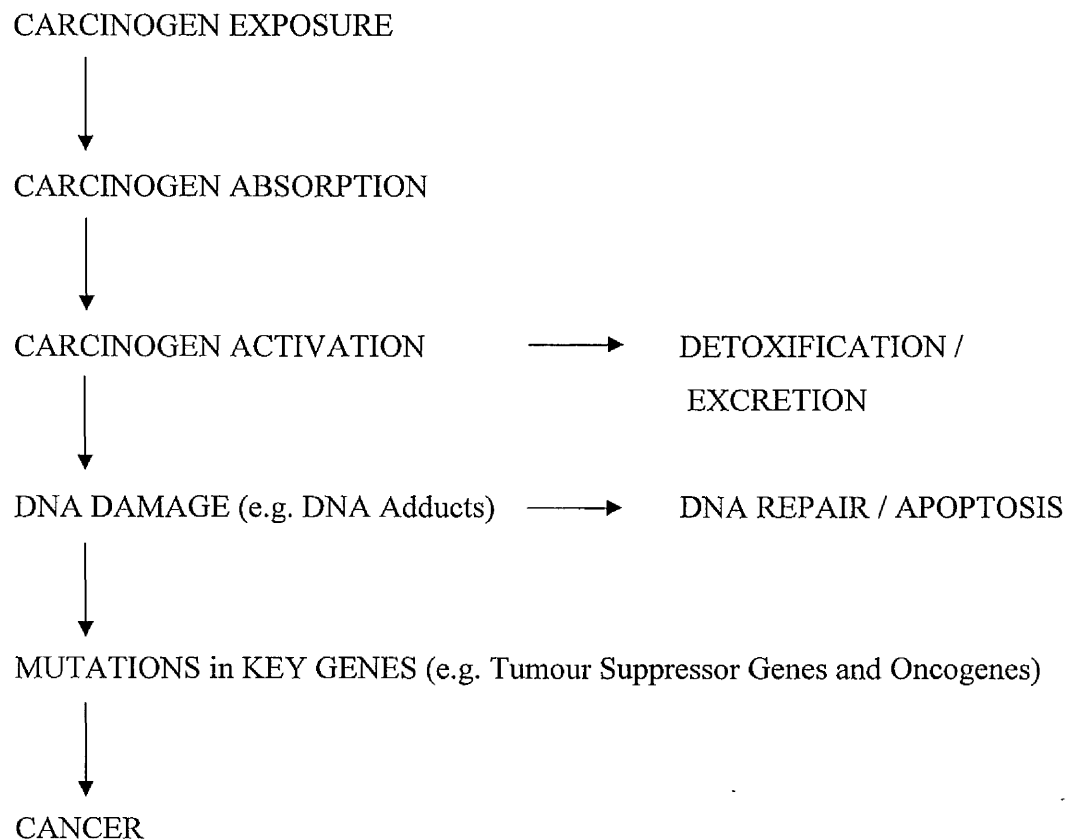
CANCER

Figure 1.9 Diagrammatic overview of the carcinogenesis of lung cancer.

1.3.1 Tobacco Smoke

Cigarette smoke is a complex aerosol made from the combustion products of tobacco. Mainstream smoke is produced at high temperatures up to 950°C and is inhaled directly through the cigarette. Sidestream smoke is produced during the smouldering of the cigarette at lower temperatures, up to 350°C and is the source of environmental tobacco smoke (Burns 1994). The majority of cigarette smoke (95%) is made from gases including nitrogen, oxygen and carbon dioxide; however, the particulate phase contains at least 4000 compounds (Hecht 1999). This complex mix of chemicals contains at least 80 carcinogens (Smith, Perfetti et al. 2003). The most important carcinogens in tobacco smoke with respect to lung cancer are thought to be alkylating agents such as the tobacco specific nitrosamines, reactive oxygen species and polycyclic aromatic hydrocarbons (Hecht 2002).

1.3.2 Alkylating Agents

Alkylating agents represent a structurally diverse family of compounds that are able to damage DNA through the formation of alkyl DNA adducts (Karran 1985). DNA adducts are formed at all available nitrogen and oxygen atoms in DNA bases and also at the phosphodiester linkages. S_N1-type alkylating agents e.g. N-methyl-N-nitroso urea (MNU) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) are able to alkylate both oxygens and nitrogens in nucleic acids whereas S_N2-type alkylating agents e.g. methyl methanesulfonate (MMS) mainly alkylate nitrogens (Drablos, Feyzi et al. 2004). Humans are exposed to alkylating agents from both endogenous sources e.g. nitrosation of amines and bile acids (Xiao and Samson 1993; Taverna and Sedgwick 1996; Shuker and Margison 1997) and exogenous sources e.g. diet (Hotchkiss 1987; Goldman and Shields 2003), cigarette smoke (Hecht 2002) and occupational exposures (Reh, DeBord et al. 2000). It has been suggested that endogenous exposure to certain alkylating agents exceeds that from exogenous sources (Tannenbaum 1980; Tricker 1997). The biological effects of alkylating agents, which include mutagenesis, clastogenicity and toxicity, are determined by the type and persistence of adducts produced.

1.3.2.1 Tobacco Specific Nitrosamines

A group of potent alkylating agents found only in tobacco products, formed from nicotine and related alkaloids are the tobacco specific nitrosamines. There are seven types of

tobacco specific nitrosamines; the most carcinogenic are N-nitrosonornicotine (NNN), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL). NNK is a potent and selective inducer of adenocarcinoma in the lungs of rodents. Since 1978 the concentration of NNK in tobacco smoke has risen by 72%. This change in the makeup of cigarettes, which has also seen the concentration of benzo[a]pyrene in smoke fall by 62% (since 1959), may explain the increased frequency of adenocarcinoma diagnosed in America (Hoffmann, Djordjevic et al. 1997). NNK is a stable compound that requires metabolic activation to produce a carcinogenic effect. The main metabolic pathways involved in the processing of NNK are shown in figure 1.10 (Hecht 1998).

Reduction of the carbonyl group of NNK produces NNAL, this is the predominant metabolite of NNK in human lung and liver tissue. This reaction is mediated by carbonyl reductases e.g. 11-β-hydroxysteroid dehydrogenase type 1. This is an important step in the detoxification of NNK, as NNAL undergoes glucuronosylation forming NNAL-glucoronide (NNAL-Gluc) which is readily excreted (Maser 2004). Less than 1% of NNK is metabolised through the α-hydroxylation pathway (Smith, Bend et al. 2003). Hydroxylation of the methylene carbon adjacent to the N-nitroso group produces methane diazohydroxide and keto aldehyde. Methane diazohydroxide methylates DNA resulting in the formation of adducts which include N7-methylguanine and O^6-methylguanine. This reaction is catalysed by a variety of cytochrome enzymes (CYP1A1, CYP2A6 and CYP3A13). Of these CYP2A13 has the highest efficiency for the metabolic activation of NNK. It is found in the respiratory tract with the greatest concentration in nasal mucosa followed by lung then trachea (Su, Bao et al. 2000). Hydroxylation of the NNK methyl group yields α-hydroxymethyl NNK. This spontaneously decomposes forming formaldehyde and 4-(3-pyridyl)-4-oxobutane-1-diazohydroxide which pyridyloxobutylates DNA. This reaction is catalysed by a variety of cytochrome enzymes including CYP2D6 and CYP2E1 (Su, Bao et al. 2000).

The metabolism of NNN can follow either α-hydroxylation pathway. The 2'-hydroxyNNN product spontaneously forms a pyridyloxobutyldiazohydroxide. This is identical to the product formed by methyl-hydroxylation of NNK. Its interaction with DNA and the properties of the resulting adducts are also identical to the NNK adducts formed via methyl-hydroxylation (Hecht 1999).
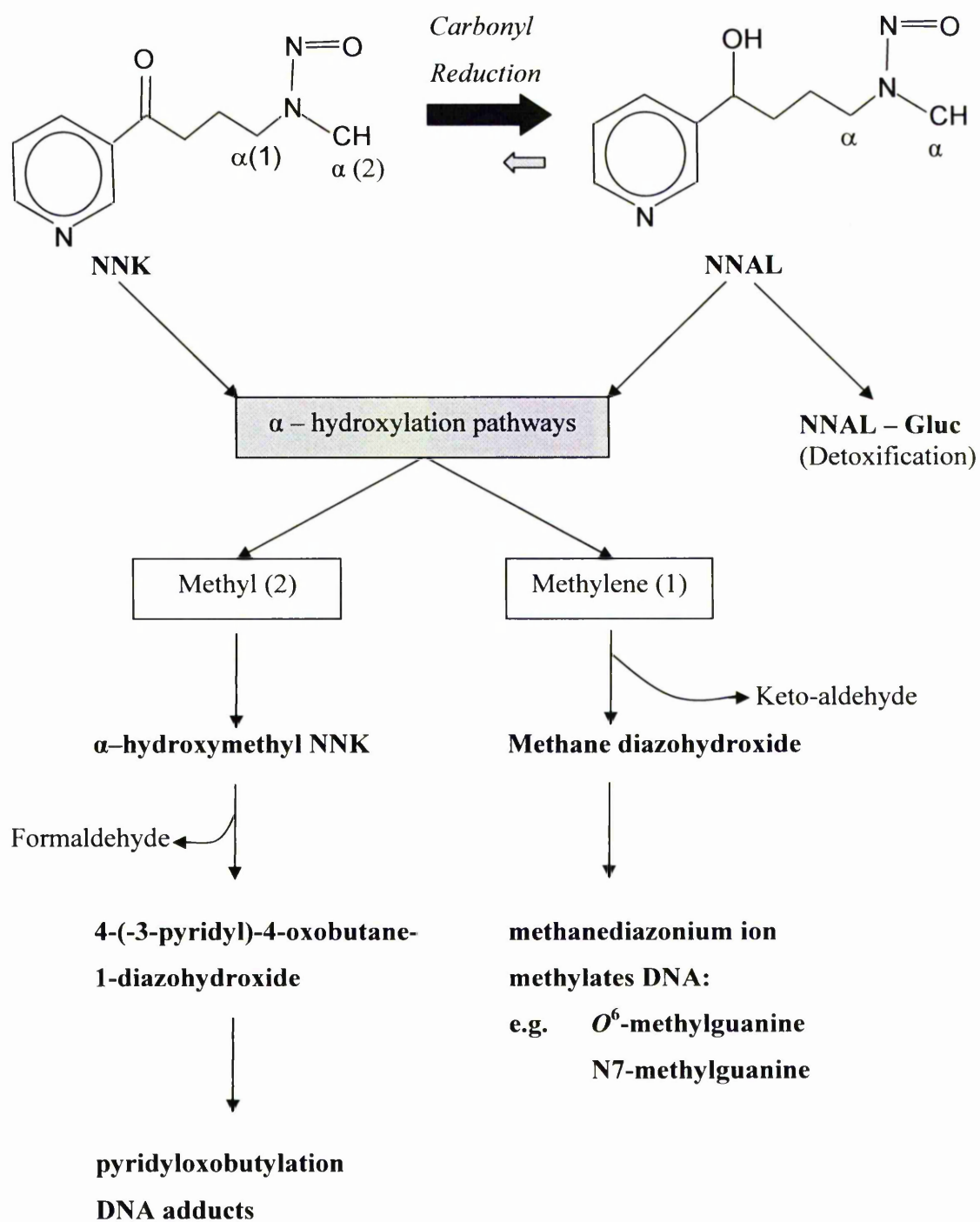
Figure 1.10 The metabolism of NNK and NNAL.

1.3.2.2 Alkylating Agent induced DNA Damage

The type and abundance of DNA lesions created by alkylating agents depends on the type of alkylating agent exposure. The most abundant lesion produced by S_N1-type alkylating agents e.g. NNK is N7-methylguanine (N7-meG) (70% all adducts), followed by methylphosphotriesters (12%), O^6-methylguanine (7%), N3-methyladenine (3%) and O^4-methylthymine (< 0.1%) (Beranek 1990; Kyrtopoulos 1998).

N7-meG is not directly mutagenic, however spontaneous or enzymatic depurination can occur to create mutagenic AP sites (Kyrtopoulos 1998; Rinne, He et al. 2005). N7-meG has been detected in bronchial specimens and peripheral blood mononuclear cells (PBMCs) from smokers and non-smokers (Shields, Povey et al. 1990; Mustonen and Hemminki 1992). Mustonen *et al* (1993) compared the levels of N7-meG in bronchial tissue and PBMCs between smokers and non-smokers. They found significantly increased adduct levels in the smoking group for both tissues compared with non-smokers and a positive correlation in adduct levels between bronchial tissue and PBMCs (r = 0.77) in smokers (Mustonen, Schoket et al. 1993). This provides evidence that smoking increases alkylation damage in DNA. A topographical study of the distribution of N7-meG within different areas of the lung showed no specific pattern (Blomeke, Greenblatt et al. 1996).

The most biologically important lesion produced by alkylating agents is O^6-meG. Increasing levels of O^6-meG have been shown to be associated with a linear increase in tumour incidence in Clara cells of the bronchial epithelium in mice treated with NNK (Belinsky, Foley et al. 1990). The presence and persistence of O^6-meG has also been shown to be critical in the initiation of lung tumours in A/J mice treated with NNK (Peterson and Hecht 1991). The presence of O^6-meG has been detected in DNA from individuals with no significant alkylating agent exposure (Georgiadis, Samoli et al. 2000). The carcinogenic mechanism of O^6-meG is complex, its ability to mispair with thymine instead of cytosine during DNA replication results in an increased frequency of GC to AT point transition mutations seen after alkylating agent exposure (Margison, Povey et al. 2003). The mechanism by which O^6-meG adducts induce GC to AT transitions is shown in Figure 1.11 (Margison, Santibanez Koref et al. 2002).

methylating agent
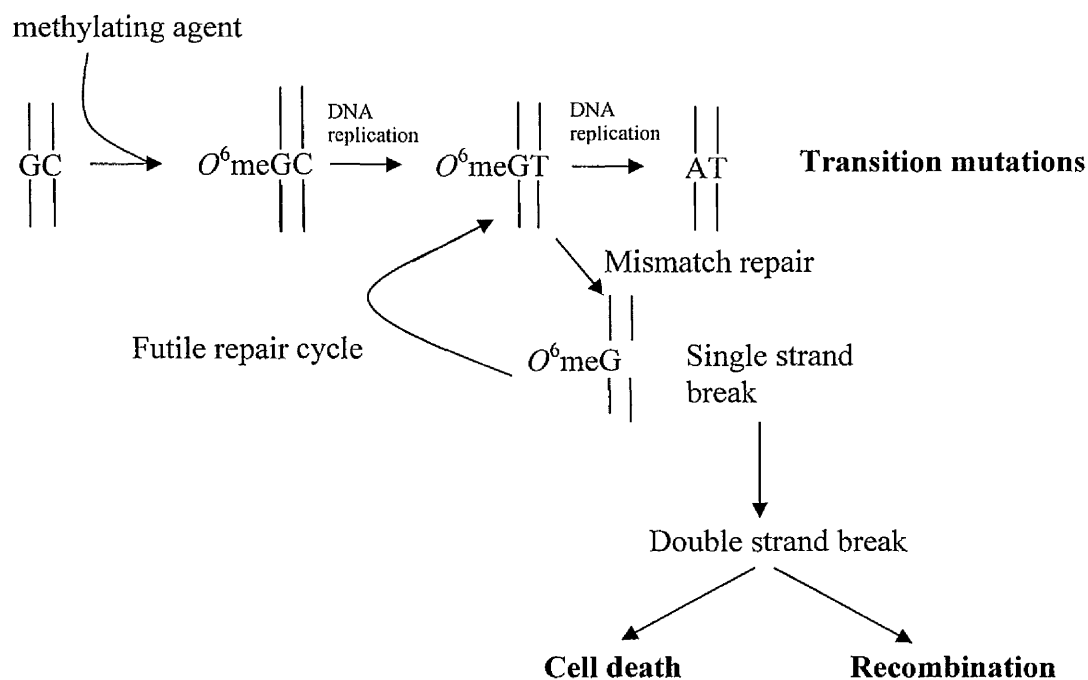
Figure 1.11 The mechanism by which O^6-meG produces biological effects of transition mutations (80%), recombination (20%) and cell death (<1%).

Pyridyloxobutyl adducts, are also thought to have an important role in NNK induced lung tumours in rats (Staretz, Foiles et al. 1997). One possible mechanism is thought to be the blocking of MGMT repair of O^6-meG resulting in sustained levels of O^6-meG in experimental models, increasing the potential for tumour induction by NNK (Peterson, Liu et al. 1993). They are also thought to induce single strand breaks in DNA which increases the risk of a mutational event (Cloutier, Drouin et al. 2001).

1.3.3 Polycyclic Aromatic Hydrocarbons

Polycyclic aromatic hydrocarbons (PAH) are found in tobacco smoke, certain foods and environmental pollution. Benzo[a]pyrene (BaP) is one of the most carcinogenic PAHs detected. BaP requires metabolic activation to produce a reactive product able to interact with cellular macromolecules. The following process metabolically activates BaP (Boysen and Hecht 2003):

1. Cytochrome P450 enzymes catalyse the formation of (7R,8R)-epoxy-7,8-dihydrobenzo[a]pyrene (BaP-7,8-oxide).

2. **BaP-7,8-oxide** is then converted to (7R,8R)-dihydroxy-7,8-dihydrobenzo[a]pyrene (BaP-7,8-diol), by the enzyme epoxide hydrolase.

3. **BaP-7,8-diol** is further oxidised, by cytochrome P450 and other enzymes, producing (7R,8S)-dihydroxy-(9R,10R)-epoxy-7,8,9,10-tetrahydro-benzo[a] pyrene (**BPDE**).

The mutagenic potential of BPDE is due to its interaction with DNA. This reaction produces a major adduct at the N2 position of deoxyguanosine (BPDE-N2-dG) (Boysen and Hecht 2003). BPDE-DNA adducts form preferentially at lung cancer mutational hotspots along the p53 gene; these adducts produce G to T transversions which if unrepaired can lead to point mutations (Smith, Denissenko et al. 2000). A matched case control study, nested in a prospective study, reported by Tang *et al* (2001), showed that healthy current smokers, who had elevated levels of aromatic DNA adduct in PBMCs at study enrolment, were three times more likely to be diagnosed with lung cancer one to thirteen years later than current smokers with lower adduct levels. This suggested that there was a role for DNA adducts in the development of lung cancer (Tang, Phillips et al. 2001).

1.3.4 Reactive Oxygen Species

Reactive oxygen species (ROS) include the hydroxyl radical ($^\bullet$OH), superoxide radical ($O_2^{\bullet-}$) and hydrogen peroxide (H_2O_2). They are created either endogenously as by-products of normal cellular metabolism, or from exogenous sources such as exposure to ultraviolet radiation (Klaunig and Kamendulis 2004). Antioxidants form an important cellular defence mechanism against the constant exposure to potentially harmful ROS. The balance between oxidants and antioxidants determines the oxidative stress of the cell. Cellular antioxidants include enzymes such as superoxide dismutase, catalase and glutathione peroxides; non-enzymic antioxidants include glutathione and vitamins C and E. Despite these defence mechanisms an estimated 10^5 oxidative lesions are formed in cells each day (Fraga, Shigenaga et al. 1990).

ROS can modify the DNA molecule directly through base or deoxyribose modifications, DNA cross-links and single or double stranded DNA breaks (Klaunig and Kamendulis 2004). Bases can also be modified prior to incorporation into DNA (Cheng, Cahill et al. 1992). There have been in excess of 100 oxidative DNA modifications identified but the lesion which has received most attention is 8-oxoguanine (8-oxoG) (Kasai 1997). 8-oxoG formed in DNA can lead to GC to TA point transversions but 8-oxoG can also be misincorporated opposite adenine during DNA replication resulting in AT to CG point transitions; this lesion is therefore potentially mutagenic (Cheng, Cahill et al. 1992). GC to TA point transversions are commonly found in p53 a tumour suppressor gene which is mutated in the majority of lung cancers (Gao, Mady et al. 2003).

The gas phase of cigarette smoke contains free radicals, the tar phase contains chemicals such as quinine and hydroquinone that can induce free radical formation (Church and Pryor 1985). Tobacco smoking is therefore able both directly and indirectly (through recruitment of macrophages) to increase the concentration of free radicals in the lungs of smokers. Evidence that smoking increases oxidative damage to DNA has been provided by studies examining levels of 8-oxoG in smokers and non-smokers. Elevated levels of 8-oxoG have been demonstrated in lung tissue (Asami, Manabe et al. 1997) and peripheral blood mononuclear cells of smokers compared to non-smokers (Kiyosawa, Suko et al. 1990; Asami, Hirano et al. 1996; Lodovici, Casalini et al. 2000). Urinary excretion of 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodG) has also been demonstrated to be higher in smokers compared to non-smokers (Pourcelot, Faure et al. 1999; Loft, Svoboda et al. 2005).

There is evidence of increased oxidative DNA damage in lung tissue (Inoue, Osaki et al. 1998) and PBMCs (Vulimiri, Wu et al. 2000) from lung cancer cases compared to controls. Studies examining matched tumour and normal lung tissue have shown either increased 8-oxoG levels in tumour tissue (Jaruga, Zastawny et al. 1994) or no difference (Hardie, Briggs et al. 2000). Interpretation of these data is hampered by the problem of temporality; raised oxidative damage to DNA may be a consequence of the presence of cancer rather than a causative factor (Olinski, Gackowski et al. 2003). Prospective studies have not been performed.

1.3.5 The Metabolic Activation and Detoxification of Tobacco Smoke Carcinogens

The majority of carcinogens contained in tobacco smoke require activation to produce their carcinogenic effects (Kiyohara, Otsu et al. 2002). Metabolic activation by phase I enzymes is balanced by detoxification undertaken by phase II enzymes (Benhamou, Lee et al. 2002). The balance between activation and detoxification could determine the levels of activated and potentially damaging carcinogens present in cells. Inherited differences in metabolic capacity of these enzymes may explain some of the variation seen in susceptibility to tobacco smoke. Many studies have looked at genetic polymorphisms, which are markers for or cause changes in metabolic enzyme activity to investigate the impact on lung cancer susceptibility of the metabolic processing of carcinogens.

1.3.5.1 Phase I Enzymes

Cytochrome P450 enzymes mediate the metabolic activation of carcinogens by catalysing their oxidation into reactive electrophilic metabolites (Raunio, Husgafvel-Pursiainen et al. 1995). These chemical intermediaries can react with DNA to produce adducts and so begin the process of carcinogenesis. Polymorphisms of the two main CYP450 enzymes involved in the metabolic activation of NNK are discussed below.

1.3.5.1.1 CYP2A6

CYP2A6 is an important hepatic cytochrome enzyme representing 1-10% of the total liver cytochrome content; however only small amounts are found in bronchial epithelial cells. It is specifically responsible for the inactivation of nicotine to produce cotinine (Raunio, Rautio et al. 2001). One study using methoxsalen to inhibit CYP2A6 activity reduced significantly tumour frequency in A/J mice treated with NNK (Takeuchi, Saoo et al. 2003; Takeuchi, Saoo et al. 2006). There are multiple allelic variants of CYP2A6 and polymorphisms are largely responsible for the variations in activity of the enzyme. Two alleles CYP2A6*2 and CYP2A6*4 have no enzymatic activity. There is a 1000 fold difference in metabolic activity between the wildtype allele and homozygotes for the inactive alleles (Xu, Goodz et al. 2002). The CYP2A6*4 polymorphism has been associated with decreased lung cancer risk in two Japanese populations (Kamataki, Nunoya et al. 1999; Ariyoshi, Miyamoto et al. 2002; Kamataki, Fujieda et al. 2005). However, a

study of lung cancer subjects in a Caucasian population showed no modification of lung cancer risk with the presence of *2 or *4 alleles, OR 1.1 (95% confidence interval 0.7-1.9) (Loriot, Rebuissou et al. 2001); findings were also mixed in two Chinese populations (Tan, Chen et al. 2001; Wang, Tan et al. 2003). Variation in risk between different populations might be due to differing polymorphism frequencies between ethnic groups. In Japanese populations *4 allele frequency is 20-31% compared to 0.5-4.9% in Caucasian populations. The *2 allele has been detected in 1.1-3.0% of Caucasians studied but never detected in Japanese populations (Xu, Goodz et al. 2002).

1.3.5.1.2   CYP2A13

CYP2A13 is the most efficient CYP450 enzyme involved in the metabolic activation of NNK. A genetic polymorphism, which is present in 2% of Caucasians, has been identified that reduces the activity of this enzyme by half (Arg257Cys amino acid substitution) (Cauffiez, Lo-Guidice et al. 2004). The presence of at least one allele with the 257Cys substitution significantly reduced the risk of adenocarcinoma in a Chinese population (OR 0.41, 95% confidence interval 0.23-0.71). A second polymorphism which produces a non-functioning protein (Arg101Stop, CYP2A13*7) was detected in 5% of a French Caucasian population. There was no significant effect on lung cancer risk in the study as a whole; however, the presence of the null allele did increase the risk of small cell lung cancer after stratification for histological type (Cauffiez, Lo-Guidice et al. 2004).

1.3.5.2 Phase II Detoxification

Glutathione S-Transferases (GST) are a multigene family of enzymes which catalyse the attachment of glutathione to reactive electrophilic compounds e.g. polycyclic aromatic hydrocarbons and products of oxidative stress. This reduces the reactivity of these chemicals with cellular macromolecules and facilitates their excretion from the body (Eaton and Bammler 1999; Parl 2005). GSTs therefore play a central role in cellular detoxification. The GSTs are divided into at least seven distinct classes (Landi 2000) and polymorphisms affecting GSTM1, T1 and P1 have been most extensively studied with respect to cancer.

1.3.5.2.1   GSTM1

The gene GSTM1 is a member of GSTM, a subfamily of GSTs, encoded by a 100kb gene cluster on chromosome 1p13. The gene cluster is arranged as 5'-GSTM4-GSTM2-GSTM1-GSTM5-GSTM3-3'; with the GSTM1 gene flanked by two identical 4.2kb regions. Homologous recombination of the flanking regions results in a 16kb deletion involving the whole gene and the creation of the GSTM1 null allele (Xu, Wang et al. 1998). A homozygous deletion of the GSTM1 gene (null genotype) affects approximately 50% of Caucasian populations (Benhamou, Lee et al. 2002; Weiserbs, Jacobson et al. 2003). GSTM1 detoxifies the reactive intermediaries of several tobacco smoke carcinogens, including BPDE formed from the metabolism of benzo[a]pyrenes (Section 1.3.3) (Nazar-Stewart, Vaughan et al. 2003; Hayes, Flanagan et al. 2005) and is expressed most strongly in the liver and to a lesser extent in the lung (Hayes and Strange 2000). Smokers with the null deletion have increased activation of carcinogens such as PAHs and reduced efficacy of detoxification in lung tissue (Nakajima, Elovaara et al. 1995). Increased activated carcinogen load is thought to increase DNA damage through the formation of DNA adducts. Higher benzo[a]pyrene diol-epoxide DNA adducts have been detected in normal lung tissue and white blood cells of current smokers who are GSTM1 null compared to those with a copy of the GSTM1 gene (Kato, Bowman et al. 1995; Rojas, Alexandrov et al. 1998; Rojas, Cascorbi et al. 2000; Alexandrov, Cascorbi et al. 2002). The association of the GSTM1 homozygous null polymorphism and lung cancer has been investigated in numerous studies. A recent meta-analysis of 119 studies (involving 19 729 lung cancer cases and 25 931 controls) concluded that the relative risk of lung cancer of the GSTM1 null genotype was 1.18 (95% CI 1.14 – 1.23) (Ye, Song et al. 2006). However the authors noted the possibility of publication bias; when the analysis was restricted to five studies containing greater than 500 cases the relative risk was lower at 1.04 (95% CI 0.95 – 1.14). The initial conclusion was consistent with an earlier meta-analysis of 43 case-control studies which found a weak but statistically significant association for the null genotype and lung cancer (OR 1.17, 95% CI 1.07-1.27) (Benhamou, Lee et al. 2002); however, further breakdown of this figure into ethnic groups revealed a lower odds ratio in Caucasians of 1.10 (95% CI 1.01-1.19) and a higher odds ratio in Asian populations of 1.33 (95% CI 1.06-1.67). A previous study in the Wythenshawe population found the null genotype to be associated with a reduced incidence of lung cancer (OR 0.50, 95% CI 0.27-0.94), this is a surprising finding considering the results of both meta-analyses discussed above (Lewis, Cherry et al. 2002).

1.3.5.2.2   GSTP1

GSTP1 is the most abundant of the GST enzymes in lung tissue (Anttila, Hirvonen et al. 1993) and has the highest specific activity towards BPDE, the active carcinogen formed from the phase 1 activation of benzo[a]pyrene (Nazar-Stewart, Vaughan et al. 2003). Two genetic polymorphisms have been studied. The first in exon 5 (Ile105Val) and the second in exon 6 (Ala114Val). The amino acid changes are in the electrophile binding site of the GSTP1 enzyme and are associated with reduced catalytic activity (Ali-Osman, Akande et al. 1997). A recent meta-analysis of both polymorphisms did not show a significant association with lung cancer susceptibility (Ye, Song et al. 2006).

1.3.5.2.3   GSTT1

The GSTT1 gene is located on chromosome 22q11.2 and is 8.1kb in length. Distinct from other GST enzymes GSTT1 is conserved among evolutionary distant species (Hiltonen, Clarke et al. 1996). GSTT1 is involved in the metabolism of methylating agents, pesticides and industrial solvents and therefore may play a role in lung carcinogenesis (Landi 2000; Raimondi, Paracchini et al. 2006). The GSTT1 positive phenotype may also increase the mutagenicity of dichloromethane (Sherratt, Pulford et al. 1997) and remove the protective effects of isothiocyanates (London, Yuan et al. 2000; Brennan, Hsu et al. 2005). The presence or absence of the GSTT1 gene did not effect the levels of benzo[a]pyrene diol-epoxide DNA adducts in one study (Rojas, Cascorbi et al. 2000). A homozygous null deletion of the GSTT1 gene is found in approximately 20% of Caucasian and 60% of Asian populations (Nelson, Wiencke et al. 1995; Raimondi, Paracchini et al. 2006). A recent meta-analysis showed the null genotype to be associated with a small increased risk of lung cancer (RR 1.13, 95% CI 1.02-1.26), there was no evidence for publication bias. However, limiting the analysis to studies with at least 500 cases showed no increased risk for the null genotype (RR 0.99, 95% CI 0.86-1.11) (Ye, Song et al. 2006). This meta-analysis included studies from different ethnic populations. A meta and pooled analysis by Raimondi *et al*, showed a significant increased risk of lung cancer for the null genotype in Asian populations but no association in Caucasian populations (Raimondi, Paracchini et al. 2006). However, in the same analysis, the GSTT1 null genotype was found to be significantly protective of lung cancer in individuals reporting occupational exposure and specifically, asbestos exposure.

1.4   DNA Repair and Lung Cancer

DNA damage is formed through three main mechanisms, firstly through exposure to environmental agents, such as carcinogens found in tobacco smoke or ultra-violet light. Secondly from exposure to endogenous agents namely products formed by normal cellular metabolism e.g. reactive oxygen species. Finally the DNA molecule can undergo spontaneous deamination under physiological conditions. For example deamination of cytosine results in the formation of uracil and adenine converts to hypoxanthine, both lesions are miscoding (Hoeijmakers 2001). The ability to repair DNA is therefore central to maintaining the integrity of the genome. The main forms of DNA repair are: damage reversal, nucleotide excision repair, base excision repair, recombination repair and mismatch repair. Inherited deficiencies in DNA repair have been associated with some forms of cancer but not lung cancer. Examples include defects in the nucleotide excision repair pathway and skin cancer (xeroderma pigmentosa), defects in mismatch repair and hereditary non-polyposis colorectal cancer and defective homologous recombination and breast and ovarian cancer (BRCA1/2) (Hoeijmakers 2001). The main pathways concerned with the repair of oxidative and alkylation damage will be discussed below.

1.4.1   Damage Reversal

The repair of the mutagenic lesion O^6-meG is undertaken by the DNA repair protein O^6-alkylguanine DNA alkyltransferase (MGMT). The O^6-methyl group is transferred to the protein's active site, a cysteine residue (codon 145). This inactivates the protein which is then ubiquitinated and degraded by the proteosomal system. MGMT also acts on other groups such as 2-chloroethyl, benzyl and pyridyloxobutyl adducts (Pegg 2000). Experiments using mice models have shown MGMT to be a critical element in the defence against the carcinogenic effects of alkylating agents. MGMT null mice although phenotypically normal were significantly more susceptible to alkylating agent toxicity and tumour induction than wildtype mice (Tsuzuki, Sakumi et al. 1996; Iwakuma, Sakumi et al. 1997; Sakumi, Shiraishi et al. 1997). The overexpression of MGMT also protected against carcinogenesis (Zhou, Manguino et al. 2001), including NNK induced lung carcinogenesis (Liu, Qin et al. 1999). The level of MGMT activity might therefore modify an individual's susceptibility to the mutagenic and carcinogenic effects of alkylating agents. Indeed, one report using cultured fibroblasts suggested an increase risk of lung cancer with low MGMT activity (Rudiger, Schwartz et al. 1989). MGMT expression is

silenced in a significant number of primary human tumours (Citron, Schoenhaus et al. 1993; Esteller, Hamilton et al. 1999). The mechanism of gene silencing is the hypermethylation of the CpG island in the promoter region of the gene resulting in loss of transcription (Danam, Qian et al. 1999). This abnormal methylation pattern has been detected in between 21-38% of NSCLC tumour tissue (Esteller, Sanchez-Cespedes et al. 1999; Wolf, Hu et al. 2001; Zochbauer-Muller, Fong et al. 2001; Brabender, Usadel et al. 2003). The proportion of tumour tissue methylated in this way is greater in adenocarcinoma (45%) than squamous cell carcinoma (23%) of the lung (Wolf, Hu et al. 2001). Pulling *et al* (2003) showed a prevalence of methylation of 51% in tumour tissue from adenocarcinoma; the prevalence was significantly increased in non-smokers compared with smokers (66% vs 47%) (Pulling, Divine et al. 2003). In a study by Wolf *et al* (2001), of 92 resected tumours p53 mutations were detected in 60% of the samples. Inactivation of the MGMT gene was present in 64% of tumours with a G to A transition mutation in the tumour suppressor gene p53; this was significantly more than in those tumours with other p53 mutations (27%, $p = 0.02$), or tumours with wild type p53 (18%, $p = 0.006$). G to A transitions were found in only 8% of unmethylated tumours. The authors suggest that MGMT protects against G to A mutations in p53, with loss of MGMT contributing to a large fraction of p53 mutations (Wolf, Hu et al. 2001). This provides evidence that MGMT inactivation may be an early change in carcinogenesis.

MGMT activity was shown not to be associated with lung cancer when examined in PBMCs in two case-control studies (Boffetta, Nyberg et al. 2002; Margison, Heighway et al. 2005). One reason for this maybe the marked variability seen in measured MGMT activity both within the same individual and between different individuals, with intra exceeding inter-individual variability (Grafstrom, Pegg et al. 1984; Gerson, Trey et al. 1986; O'Donnell, Barber et al. 1999; Margison, Povey et al. 2003; Povey, O'Donnell et al. 2006). A 3.5 fold variability in MGMT activity in PBMCs has also been seen in samples taken from the same individual over several weeks (Janssen, Eichhorn-Grombacher et al. 2001) and circadian variability from samples taken over a 24 hour period (30%) (Marchenay, Cellarier et al. 2001). The effect of smoking on MGMT activity has been investigated in human tissues with mixed results. Increased activity has been seen in normal lung tissue (Drin, Schoket et al. 1994), tumour tissue (using immunohistochemistry) (Mattern, Koomagi et al. 1998) and placental tissue (Slupphaug, Lettrem et al. 1992) from smokers compared to non-smokers. Smoking had no effect in PBMCs (Vahakangas, Trivers et al. 1991; Hall, Bresil et al. 1993; Boffetta, Nyberg et al.

2002; Margison, Heighway et al. 2005; Povey, O'Donnell et al. 2006), normal lung tissue (Nakatsuru, Tsuchiya et al. 1994) or bronchoalveolar cells (Vahakangas, Trivers et al. 1991) in further studies. One study has reported a decrease in activity in bronchial epithelial cells in smokers compared to non-smokers (Povey, O'Donnell et al. 2006).

Polymorphic variation in the MGMT protein and lung cancer risk has been investigated in a limited number of studies. A polymorphism in codon 143 which involves an isoleucine (ATC) to valine (GTC) substitution has been studied because of the close proximity it holds to the central cysteine acceptor residue at codon 145 (Deng, Xie et al. 1999). A polymorphism in codon 178, involving a lysine (AAG) to arginine (AGG) substitution, has also been studied because it is in 100% linkage disequilibrium with codon 143 and it contains a restriction enzyme site (Cohet, Borel et al. 2004). Functional studies of this polymorphism have shown no difference between wildtype and variant in ability to repair O^6-meG in DNA, however wildtype protein was more easily inactivated by O^6-benzylguanine and O^6-(4-bromothenyl)guanine, both of which are undergoing clinical trials as MGMT inactivating agents (Margison, Heighway et al. 2005).

Kaur *et al* were the first to examine the codon 143 genotype with respect to lung cancer risk. The study population included 139 histologically proven primary lung cancers in a mixed ethnicity population (56 Caucasian and 83 African American) and 136 matched (age, sex, race) hospital based controls (Kaur, Travaline et al. 2000). Cases were more likely than controls to have been ever smokers. In the combined population there was a significantly increased risk of lung cancer in subjects heterozygous for I143V (OR 2.1, 95% CI = 1.01-4.7, p = 0.05) compared to wildtype homozygotes. When Caucasian subjects were analysed separately the increased odds did not reach significance (OR 2.0, 95% CI = 0.78-5.7). An increased risk for subjects heterozygous for the 143V/178R polymorphism (OR 2.05, 95% CI = 1.03-4.07) was also seen in a multi-centre study of 136 cases and 133 hospital/population controls. All study subjects were non-smokers and 87.5% of cases were female. The risk varied with histology with increased risk seen with adenocarcinoma but not squamous cell carcinoma (Cohet, Borel et al. 2004). Three further studies have found no association with lung cancer risk and the K178R polymorphism (Krzesniak, Butkiewicz et al. 2004; Yang, Coles et al. 2004; Zienolddiny, Campa et al. 2005).

1.4.2 Base Excision Repair (BER)

BER is the main repair mechanism directed against DNA damage induced by reactive oxygen species and alkylating agents. It is also responsible for the repair of apurinic / apyrimidinic (AP) sites which are formed as a result of spontaneous depurination or as intermediaries of BER itself (Hang and Singer 2003). Depurination could produce as many as 10 000 AP sites / day / cell (Lindahl and Barnes 2000) and therefore represents one of the most abundant types of DNA damage (Nakamura and Swenberg 1999). The initiation of BER is the function of lesion specific DNA glycosylases. Two types of DNA glycosylase have been described; monofunctional DNA glycosylases e.g. MPG flip the damaged base out of the DNA helix then cleave the N-C1′ glycosylic bond between the abnormal base and the deoxyribose to create an AP site. Bifunctional DNA glycosylases e.g. hOGG1 cleave the 3′ side of the AP site after glycosylase activity by means of an additional AP lyase function (Hang and Singer 2003).

The resulting AP site is processed down one of two BER pathways, either a short patch or long patch repair pathway. In short patch repair, the process of repairing a single base gap is started by apurinic/apyrimidinic endonuclease-1 (APE-1), which cleaves the phosphodiester bond 5′ to the AP site. Poly (ADP-ribose) polymerase I (PARP1) is activated by and binds to single strand breaks and is thought to play a role in BER possibly through recruitment of components of the pathway (Barnes and Lindahl 2004; Curtin 2005). DNA polymerase beta (Polβ) removes the 5′ terminal deoxyribose phosphate residue before inserting the appropriate single nucleotide, then DNA ligase I or the DNA ligase III/XRCC1 complex seal the repaired patch. In long patch repair, a gap of up to 10 nucleotides is filled by Pol δ or ε; the process is dependent on proliferating cell nuclear antigen (PCNA) and the resultant flap is removed by the FEN1 endonuclease (Hang and Singer 2003). Inherited deficiencies in BER have not been identified in humans. The inactivation of BER core proteins results in embryonic lethality, underlining the fundamental importance of this repair mechanism (Hoeijmakers 2001).

1.4.2.1 8-Oxoguanine DNA N-Glycosylase (hOGG1)

8-Oxoguanine DNA N-glycosylase (hOGG1) is a bifunctional DNA glycosylase which repairs the mutagenic lesion 8-oxoguanine (8-oxoG) when paired with cytosine in double stranded DNA, via the short patch BER pathway (Figure 1.12). The hOGG1 gene is

located on chromosome 3p25. Alternative splicing produces two forms, an α form which is exclusively found in the nucleus and is more abundant than the β form which is found on the inner membrane of mitochondria (Boiteux and Radicella 2000). hOGG1, in co-operation with at least two other genes, provides a defence barrier against the mutagenic potential of 8-oxoG, the additional genes are MYH and MTH. MYH is a monofunctional glycosylase that removes adenine that has been misincorporated opposite 8-oxoG in double stranded DNA to restore a 8-oxoG:C pair which is then available to be repaired by hOGG1. MTH hydrolyses 8-oxoGTP in nucleotide pools preventing its incorporation into DNA (Cooke, Evans et al. 2003). The importance of the hOGG1 gene in relation to lung cancer development has been investigated in a number of different epidemiological and molecular biological studies. The effect of homozygous OGG1 gene deletion (OGG1$^{-/-}$) has been studied in mice models with mixed results. Klungland *et al*, reported no overt phenotype in the knockout mice; there was no increase in tumour formation despite an elevation of 8-oxoG and a moderate increase in spontaneous mutation rate (Klungland, Rosewell et al. 1999). Arai *et al*, also showed no increase in tumour formation in OGG1$^{-/-}$ mice chronically exposed to the oxidising agent potassium bromate (Arai, Kelly et al. 2006). Osterod *et al*, showed increased 8-oxoG levels were restricted to slowly proliferating tissues with high oxygen metabolism e.g. liver, in OGG1$^{-/-}$ mice suggesting the presence of a back up mechanism for repair of 8-oxoG (Osterod, Hollenbach et al. 2001). However Sakumi *et al*, showed increased tumour formation in OGG1$^{-/-}$ mice (Sakumi, Tominaga et al. 2003).

Tumour development in double knockout (MYH$^{-/-}$, OGG1$^{-/-}$) mice has been a more consistent finding. Two thirds of double knockout mice developed tumours predominantly of the lung and ovary and also lymphomas in one study; life expectancy was also significantly reduced (Xie, Yang et al. 2004). A second group also reported increased spontaneous tumour formation in lung and small intestine of double knockout mice (MYH$^{-/-}$, OGG1$^{-/-}$) without exogenous oxidant exposure (Russo, De Luca et al. 2004). The difference in phenotype between single and double knockout mice suggests that the oxidative repair mechanisms of the cell can cope with the loss of one gene but not two. Interestingly three quarters of tumours seen in the double knockout mice contained GC:TA point mutations in codon 12 of the *K-ras* oncogene, no mutations were seen in normal tissue. Activation of the codon *K-ras* oncogene is a common finding in NSCLC especially adenocarcinomas (Jacobson 1999). This suggests a causal role of oxidative DNA damage in carcinogenesis.
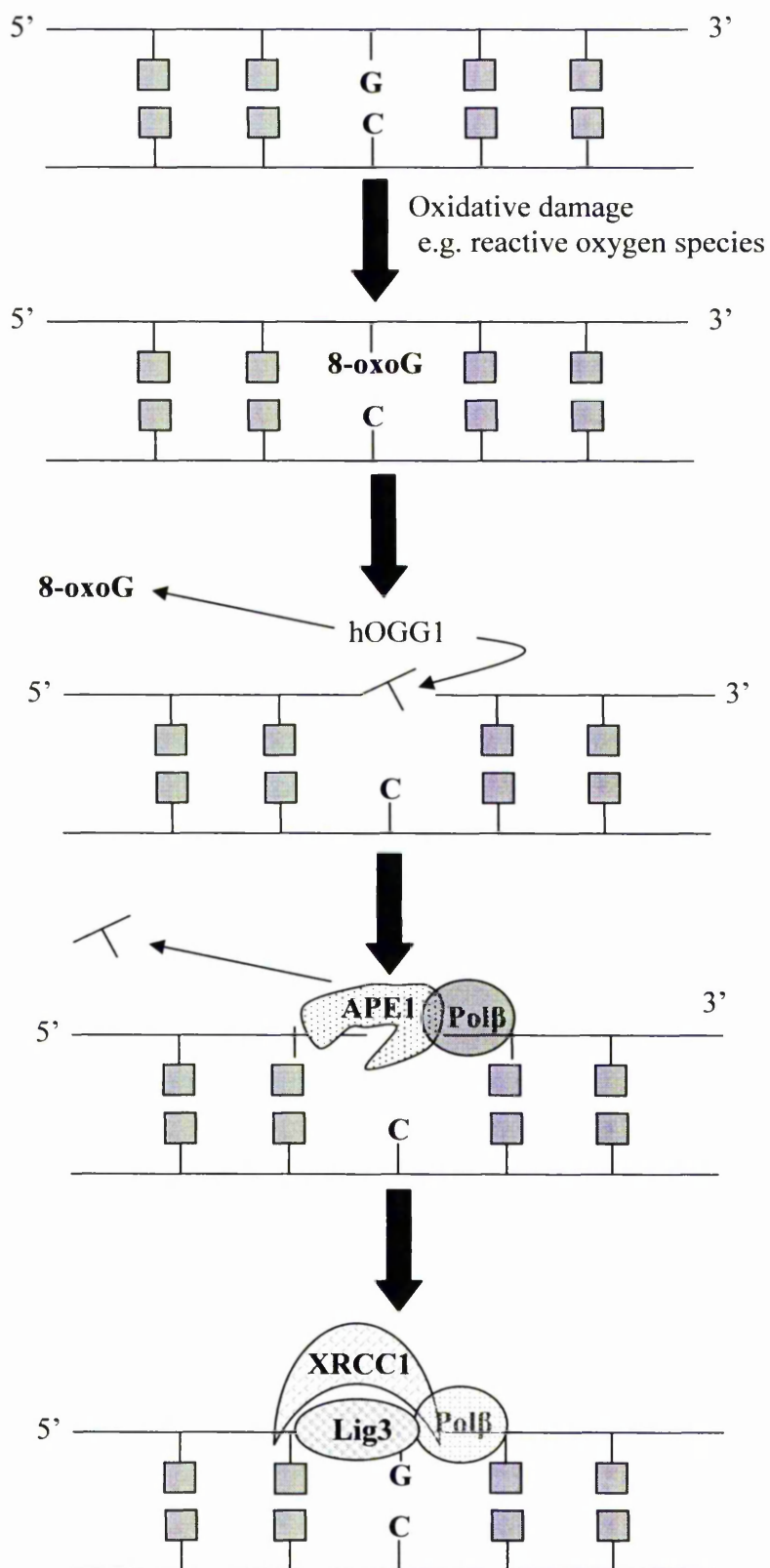
Figure 1.12 hOGG1 initiated repair of 8-oxoG (an example of base excision repair).

Most epidemiological studies investigating the role of hOGG1 in lung cancer development have focused on a single nucleotide polymorphism (C to G transversion at nucleotide 1245, rs1052133) which results in a serine to cysteine substitution in codon 326. This polymorphism is present in 22% of Caucasians and 42% of Japanese (Le Marchand, Donlon et al. 2002). Studies investigating the functional effect of this polymorphism have produced inconsistent findings (Weiss, Goode et al. 2005). This inconsistency is also seen in the results of epidemiological studies. Park *et al*, showed an increased risk of developing lung cancer for 326Cys/Cys homozygotes (OR 3.8, 95% confidence interval 1.4-10.6) and 326Cys/Ser heterozygotes (OR 1.9, 95% confidence interval 1.2-2.9) (Park, Chen et al. 2004). This finding was supported by two further studies using Caucasian populations but only in Cys/Cys homozygotes (Le Marchand, Donlon et al. 2002; Zienolddiny, Campa et al. 2005). A study in a Japanese population found the polymorphism to be a risk for adenocarcinoma of the lung (Kohno, Kunitoh et al. 2006). The study by Le Marchand *et al* consisted of a mixed ethnicity study population; when the Caucasian population was analysed no association with the hOGG1 polymorphism was seen with lung cancer. The study by Zienolddiny *et al* may not be generalisable as cases consisted purely of patients with operable disease; this would account for only a small proportion of lung cancers in the general population. Several other studies have found no association between the Ser326Cys polymorphism in either heterozygotes or homozygotes and lung cancer risk (Wikman, Risch et al. 2000; Vogel, Nexo et al. 2004; Hung, Brennan et al. 2005; Sorensen, Raaschou-Nielsen et al. 2006). The largest of these negative studies, Hung *et al*, was a multicentre study with in excess of 4000 participants.

Functional enzyme studies have reported that hOGG1 repair activity is significantly reduced in lymphocytes from lung cancer cases compared to cancer free controls (Gackowski, Speina et al. 2003; Paz-Elizur, Krupsky et al. 2003).

1.4.2.2 3-Methylpurine DNA glycosylase (MPG)

3-methyladenine lesions, produced after endogenous or exogenous exposure to alkylating agents (Marnett and Burcham 1993), are able to block DNA synthesis and are therefore cytotoxic (Larson, Sahm et al. 1985; Varadarajan, Shah et al. 2003). A DNA glycosylase that is able to excise 3-meA from DNA has been found in bacteria, yeast, plants, rodents and humans (Wyatt, Allan et al. 1999). MPG is a monofunctional DNA glycosylase located on chromosome 11. In addition to 3-meA, it has a very broad range of substrates

(Table 1.3), broader than most other DNA glycosylases (Singer and Hang 1997). It is interesting to note that the MPG enzyme in humans and *Saccharomyces cerevisiae* is able to remove normal guanines, in *E.coli* this ability is extended to all guanines, adenines, cytosines and thymines (Berdal, Johansen et al. 1998; O'Brien and Ellenberger 2003; O'Brien and Ellenberger 2004).

Table 1.3 Known substrates released by MPG in humans and other organisms.

Known substrates released by MPG (Wyatt, Allan et al. 1999)	
In Humans	**Additional lesions in other organisms**
3-methyladenine	O^2-methylthymine
7-methylguanine	O^2-methylcytosine
3-methylguanine	7-chloroethylguanine
7-(hydroxymethyl)/ 1,N^6-ethenoadenine	7-hydroxyethylguanine
3,N^2-ethenoguanine	3-chloroethylguanine
Deaminated adenine (hypoxanthine)	5-hydroxymethyluracil
Normal Guanine	5-formyluracil
	Normal Adenine, Thymine and Cytosine

There is evidence that MPG protects against the toxic, clastogenic and mutagenic effects of alkylating agents and oxidising agents. In murine models, MPG has been shown to protect against the cell killing effects of the alkylating agents methyl methanesulfonate (MMS), 1,3-bis (2-chloroethyl)-1-nitrosurea (BCNU) and mitomycin C (MMC) (Engelward, Dreslin et al. 1996). MMS induced 3-meA lesions have been shown to inhibit DNA replication in a MPG deficient murine cellular model and prevent cells progressing through S phase (Engelward, Allan et al. 1998). Toxicity may be induced by activating programmed cell death. 3-meA and other lesions, including DNA interstrand cross links, produced by the alkylating agents MMC and BCNU appear to induce apoptosis in the same cellular murine model. This effect was increased in MPG null cells (Allan, Engelward et al. 1998; Engelward, Allan et al. 1998). MPG activity may therefore protect against some of the effects of alkylating chemotherapeutic agents.

MPG is able to excise endogenously produced mutagenic lesions including hypoxanthine and 1,N^6-ethenoadenine. However the loss of MPG did not increase the spontaneous mutation rate in *E.coli* or *S.cerevisiae* (Xiao and Samson 1993; Sidorkina, Saparbaev et al.

1997). εA can be produced exogenously by the metabolic products of the liver carcinogens vinyl chloride and ethyl carbamate or endogenously by the reaction of lipid peroxidation products with DNA (el Ghissassi, Barbin et al. 1995; Chung, Chen et al. 1996). εA is a mutagenic lesion which has been shown to induce AT to TA transversions in mammalian cells (Pandya and Moriya 1996). MPG is the only enzyme known to release εA from DNA (Saparbaev, Kleibl et al. 1995; Pandya and Moriya 1996; Hang, Singer et al. 1997). However, although MPG null mice have been shown to have higher levels of εA after treatment with vinyl carbamate, no increase in carcinogenesis was detected (Barbin, Wang et al. 2003).

Murine models, which express no MPG activity (MPG -/-), result in outwardly normal, fertile animals with a normal lifespan (Engelward, Weeda et al. 1997; Elder, Jansen et al. 1998) and no increase in spontaneous tumour induction (Parsons and Elder 2003). Assays of various tissues from these mice showed MPG to be the major DNA glycosylase in the repair of εA, 3-meA and N7-meG (Engelward, Weeda et al. 1997; Hang, Singer et al. 1997). However a more recent paper suggests that MPG may not be the most efficient repair pathway for N7-meG (Smith and Engelward 2000), with nucleotide excision repair being an alternate repair pathway (Plosky, Samson et al. 2002). After treatment with methylating agents, increased persistence of N7-meG in the liver of null mice was noted compared with normals. MMS treatment resulted in up to fourfold increase in hypoxanthine guanine phosphoribosyltransferase gene mutations, mainly AT to TA transversions, in splenic T lymphocytes in null compared to normal mice. The authors suggested this showed an important role for MPG in preventing the mutagenic effects of N-alkylpurines (Elder, Jansen et al. 1998).

A second group using MPG knockout mice reported that MPG null embryonic stem cells were more sensitive to cell death than heterozygous or wildtype cells exposed to simple methylating agents (MMS, MeOSO$_2$Et-lexitropsin) (Engelward, Dreslin et al. 1996; Engelward, Weeda et al. 1997). Interestingly heterozygous cells were no more sensitive to these agents than wildtype cells (Engelward, Dreslin et al. 1996). MeOSO$_2$Et-lexitropsin specifically induces 3-meA giving direct evidence that this lesion is cytotoxic. MPG null cells were also more susceptible to alkylation (MMS, BCNU) induced sister chromatid exchange than heterozygotes or wildtype. MPG therefore protects against cell death and chromosome damage induced by alkylating agents.

MPG activity in humans has previously been examined in only one study (Hall, Bresil et al. 1993). MPG activity was assayed along with MGMT in PBMCs from 20 current smokers and 17 non-smokers. MPG activity was higher in smokers compared to non-smokers and showed a bimodal distribution when all subjects or smokers alone were examined. MPG and MGMT activities were not correlated. There is only limited understanding regarding the expression and regulation of the MPG gene. MPG gene expression is cell cycle dependent (Bouziane, Miao et al. 2000) and has been shown to be inducible by exposure to alkylating agents in rodents (Hall, Bresil et al. 1990).

MPG activity has never been examined with respect to cancer in humans. Previous studies examining MPG mRNA expression have reported higher mRNA levels in tumours of the ovary, thymus and brain (astrocytic) when compared to normal tissue (Kim, Lee et al. 1998; Kim, An et al. 2002; Kim, Ahn et al. 2003). Greater MPG expression has also been reported in breast cancer cell lines when compared to normal and immortalised epithelial breast cells; this was not due to gene amplification and the authors suggested a causal link with breast cancer development (Cerda, Turk et al. 1998). Gene expression has been reported to correlate with MPG protein levels.

Several studies have examined the relationship between MPG overexpression and resistance to alkylating agent induced DNA damage. Although one study by Ibeanu *et al*, found that overexpression had no effect, the majority of studies report that MPG overexpression sensitises cells to the damaging effects of alkylating agents (Coquerelle, Dosch et al. 1995; Calleja, Jansen et al. 1999; Fishel, Seo et al. 2003; Rinne, Caldwell et al. 2004; Trivedi, Almeida et al. 2005). Deleterious effects of alkylation damage included increased chromosomal aberrations, stronger inhibition of DNA replication and increased DNA breaks (Coquerelle, Dosch et al. 1995).

One possible explanation for the deleterious effects of MPG overexpression on cells could be the transformation of N7-meG lesions, the most common lesion produced by methylating agents, into toxic intermediaries of the BER pathway e.g. AP sites. Evidence for this comes from two recent studies. Connor *et al*, measured survival, mutation induction and double strand break formation in *Saccharomyces cerevisiae*, which lack endogenous MPG, expressing either wildtype, inactive or catalytic variants of human MPG (Connor, Wilson et al. 2005). Using the alkylating agent MMS (adduct burden 70% N7-MeG, 10% 3-MeA (Beranek 1990)), they showed that *S. cerevisiae* expressing inactive

MPG were highly sensitive to alkylation damage. *S. cerevisiae* expressing a MPG variant that was only able to repair 3-meA and not N7-meG were more resistant to alkylation damage than the inactive or wild type protein. The authors suggested it was the ability of wild type protein to repair N7-meG that contributed to the toxicity of the model.

A study by Rinne *et al*, showed that breast cancer cells overexpressing wildtype MPG protein were more sensitive to the alkylating agents MMS and temozolomide than cells expressing inactive MPG protein or controls (Rinne, He et al. 2005). However exposure to the alkylating agent $MeOSO_2(CH_2)_2$-lexitropsin, which specifically produces 3-meA lesions (>99%), showed no differential effect on any of the cells. MMS treatment induced large numbers of N7-meG lesions in controls cells. In contrast, very few N7-meG lesions were detected in cells that overexpressed MPG. MPG overexpression did not affect the cytotoxicity induced when several other alkylating agents were used. The cytotoxicity of MPG overexpressing cells was therefore thought to be due specifically to the DNA glycosylase activity of the enzyme rapidly repairing N7-meG resulting in an increase in toxic BER pathway intermediaries. The authors had shown in a previous study that cells overexpressing MPG treated with MMS did indeed develop increased levels of BER pathway intermediaries e.g. AP sites and single strand breaks (Rinne, Caldwell et al. 2004).

1.5 Aims

The overall aim of this project was to investigate genetic factors that might modify susceptibility to the development of lung cancer or biomarkers that might indicate increased risk of lung cancer. The principal constituent of this work was a case-control study; the primary objectives were to compare incident lung cancer cases with cancer free controls for:

1. Activity of MGMT and MPG in PBMCs.
2. Distribution of genetic polymorphisms in:
   a. Phase II detoxifying enzymes
      i. GSTM1 (homozygous null genotype)
      ii. GSTT1 (homozygous null genotype)
   b. DNA repair proteins:
      i. MGMT – codon 178 and rs12268840 (intron 1)
      ii. hOGG1 – codon 326
3. GSTM1 copy number and lung cancer risk.
4. Allelic expression imbalance of MGMT in PBMCs and also in lung and tumour tissue removed from patients undergoing surgical resection of lung cancer.
5. DNA adduct levels in PBMCs.
   a. 8-oxoG
   b. N7-meG

Study subjects were interviewed using a risk factor questionnaire to collect data regarding age, gender, smoking and occupational exposure. A secondary aim was to determine the interaction between these factors and the biological measurements of exposure e.g. DNA adducts or DNA repair.

Chapter Two

Materials and Methods

2.1 Materials

2.1.1 Enzymes

Taq DNA polymerase and *Stu*1 were purchased from Promega; *Aci*1, *Pst*1, OGG1, APE-1, *BsmA*1 and *Bsa J*1 from New England Biolabs, *Ita*1 from Roche Applied Science, RNase A from Sigma and Proteinase K from Qiagen. AAG was kindly gifted by Thomas Lindahl. All enzymes were stored at -20°C.

2.1.2 Buffers

All buffers were made up with ddH_2O.

2.1.2.1 Enzyme Buffers

Storage Buffer A - 50mM tris-HCl (pH 8.0), 100mM NaCl, 0.1mM EDTA, 1mM DTT, 50% glycerol and 1% Triton® X-100

Taq DNA polymerase 10x buffer (concentrations given for 1x) - 50mM KCl, 10mM Tris-HCl (pH 9.0), 1.5mM $MgCl_2$ and 0.1% Triton® X-100

Buffer H - 50mM Tris-HCl (pH 7.5), 10mM $MgCl_2$, 100mM NaCl, 1mM DTE

Buffer Aci - 100mM NaCl, 10mM Tris-HCl (ph 7.4), 0.1mM EDTA, 1mM DTT, 200µg/ml BSA and 50% glycerol

*BsmA*1 Buffer - 300mM NaCl, 10mM Tris-HCl (pH 7.4), 0.1mM EDTA, 1mM DTT, 500µg/ml BSA and 50% glycerol

NEBuffer 2 – 50mM NaCl, 10mM Tris HCl, 10mM MgCl, 1mM DTT

2.1.2.2 Cleavage Assay Buffers

Buffer I - 50mM Tris-HCl, pH8.3, 1mM EDTA, 3mM DTT

Buffer I + leupeptin - buffer I + leupeptin 10mg/ml

Buffer IBSA - buffer I containing 1mg/ml BSA

Buffer K - 25mM HEPES-KOH, 150mM KCl, 0.5mM DTT, 0.5mM EDTA, 1% glycerol, pH 7.8

Buffer KBSA – buffer K containing 1mg/ml BSA

Cleavage Buffer - 25mM Tris-HCl, pH 7.6, 50mM KCl, 5mM EDTA

Magnesium Cleavage Buffer - cleavage buffer + 10mM $MgCl_2$

10x NEBuffer 3 (concentrations given for 1x) - 100mM NaCl, 50mM Tris-HCl, 10mM $MgCl_2$, 1mM DTT, pH 7.9

2.2   Case-Control Study Methods

A hospital based case-control study was designed to compare certain biological variables between incident lung cancer patients and cancer free controls.

2.2.1   Ethical Approval

This study was granted ethical approval by North Manchester Local Research Ethics Committee (reference 02/NM/496). Full written informed consent was given by all subjects prior to study participation.

2.2.2   Sampled Populations

Patients were recruited from two different sources within the North West Lung Centre at Wythenshawe Hospital. The first source, from which all cases and one control group were recruited, was the Bronchoscopy Unit. The second control group was recruited from the Chest Clinic.

2.2.3   Study Participation Criteria

Patients were invited to participate in the study if they met the inclusion and exclusion criteria. Differences in recruitment criteria between patients recruited in the Bronchoscopy Unit and those recruited in the Chest Clinic are detailed below (Table 2.1).

Table 2.1 Study participation criteria.

	Inclusion Criteria	Exclusion Criteria
All Subjects	aged over 40 Caucasian able to give written informed consent to study participation	previous diagnosis of lung cancer previous diagnosis of any other cancer
Bronchoscopy Unit	due to undergo bronchoscopy in the Bronchoscopy Unit	undergoing bronchoscopy for investigation of interstitial lung disease or for microbiological samples
Chest Clinic	current / former smokers or history of COPD	primary diagnosis of interstitial lung disease, asthma or bronchiectasis

2.2.4 Definition of Case and Control Status

Cases were defined as those patients who were newly diagnosed with lung cancer (i.e. incident cases). Lung cancer was definitively diagnosed if a histological sample was reported as showing lung cancer by a consultant pathologist. A clinical diagnosis was made when no histology was available but the combination of clinical history, examination, bronchoscopy findings and radiological findings were consistent with a diagnosis of lung cancer.

The controls from the Bronchoscopy Unit were defined as those patients who after investigation were found not to have lung cancer or any other malignant disease. The categorisation of patients as cases or controls was possible only after the results of all investigations e.g. bronchoscopy, CT scanning and histology were available. The Chest Clinic control group were patients with no history of malignant disease; who were or had been smokers and / or had a history of chronic obstructive pulmonary disease.

2.2.5 Study Recruitment

The study was explained to patients attending for bronchoscopy or attending the Chest Clinic. Patients who expressed an interest were given an information leaflet (Appendix 2) detailing the purpose of the study, what study participation would involve and the potential risks and benefits of taking part. Patients who gave their consent to study participation were asked to sign a consent form and were interviewed using a standard questionnaire. A blood sample was taken from all study participants.

2.2.6 Study Interview

All study participants were interviewed individually by the author. All bronchoscopy patients were interviewed prior to their procedure to ensure blinding from the results of the test. The interviews were approximately ten minutes in length and a standard risk factor questionnaire was used throughout (Appendix 1). The questionnaire was designed to collect demographic details and also information about smoking history, occupational history, occupational exposures, medication history, family and past respiratory history. It was based on a questionnaire used in previous research studies in the same population (Lewis, Cherry et al. 2002).

2.2.7 Smoking Definitions

Smokers were defined as ever smokers if they had smoked a tobacco product, such as cigarettes, cigars or a pipe, at least once a week for a year. All other subjects were classified as lifelong non-smokers. Ever smokers were subdivided into current smokers or ex-smokers if they had abstained from smoking tobacco for at least one week prior to interview. Packyears were used as a measure of reported lifetime smoking exposure and were only calculated for ever smokers of cigarettes. This measure was defined as:

duration of smoking (years) x ([average daily cigarette intake]/20)

2.2.8 Study Size Calculation and Statistical Analysis of Data

Previous studies showed the GSTM1 null genotype to be present in approximately 50% of Caucasian populations. To detect an odds ratio of 2.0 with $\acute{\alpha} = 0.05$ and $\beta = 0.20$ a sample size of approximately 150 cases and controls was required. All data collected from the interview was inputted into an SPSS database (SPSS version 13.0). Information resulting from clinical investigations and laboratory experiments was also placed in the same database. Continuous data was presented as mean values with standard deviations if normally distributed and median values with ranges if non-normally distributed. Measures of kurtosis and skewness were used to assess the normality of continuous data distribution. Differences in mean values were tested with t tests. Non-normally distributed data was log transformed to normalise distribution e.g. N7-medG. MGMT activity was expressed as geometric means with standard deviations. Categorical data e.g. genotyping was expressed as number of subjects and frequencies. Chi squared tests were used (or Fishers exact test if a value in one category was < 5) to determine significant differences in categorical data. Odds ratios were calculated by binary logistic regression analysis. Adjustments were made for age, gender and smoking duration with respect to genotyping data and lung cancer risk. Subset analysis was performed on genotyping data with the understanding that the introduction of multiple statistical testing increased the likelihood of a significant finding occurring by chance (type 1 error). One mechanism to reduce the likelihood of this occurring is known as the Bonferroni method. This method was not performed but an awareness of the risks of multiple testing was factored into the interpretation of data. Trend test were used to ascertain trends in categorical data.

2.2.9 Measures of Deprivation

The Index of Multiple Deprivation 2004 (IMD 2004), was used as a guide to deprivation of study subjects. The postcode of the subject's current address was entered into the UK National Statistics website (http://neighbourhood.statistics.gov.uk/dissemination/) which has ranked all wards in England into order of deprivation with 1 being the most deprived and 32 482 the least deprived. A mean of this value was calculated for cases and controls.

2.2.10 Biological Sample Collection and Processing

A blood sample was taken from all patients. Approximately 5ml of blood was collected and allowed to clot; a further 16ml was taken and placed in tubes containing EDTA. All samples were processed and stored in $-80^{\circ}C$ freezers on the day they were taken.

2.2.10.1 PBMC Separation

PBMCs were separated at room temperature using the separation medium Ficoll-Paque (Amersham). Whole blood (9ml) was carefully layered onto a 9ml aliquot of Ficoll-Paque. The layered blood was centrifuged at 400g for 30 minutes. This resulted in a top layer of plasma, a buffy coat of PBMCs, Ficoll-Paque medium and finally a red blood cell layer. The plasma layer was removed. The buffy coat was then gently removed and washed twice in PBS (centrifuged at 400g for 12min at $22^{\circ}C$). The PBMC cell pellet was resuspended in 1ml of PBS and immediately stored in a $-80^{\circ}C$ freezer.

2.2.10.2 The Clotted Blood Sample

The clotted blood sample was collected and immediately stored in a fridge at $4^{\circ}C$ until centrifugation on the same day (400g for 12min at $22^{\circ}C$). Serum and clotted blood were separated from the sample and immediately stored in a $-80^{\circ}C$ freezer.

2.3   Surgical Case Series Methods

The surgical case series was designed to collect lung tissue from patients undergoing surgical resection of lung cancer.

2.3.1   Ethical Approval

This study was granted full ethical approval by North Manchester Local Research Ethics Committee (reference 02/NM/495). Full written informed consent was given by all subjects prior to study participation.

2.3.2   Recruitment

Patients were recruited from the Cardiothoracic Surgical Unit at Wythenshawe Hospital. All patients attending the Unit who were due to undergo surgical resection of lung tissue for the treatment of lung cancer were eligible to participate. Patients were excluded if there was a previous diagnosis of any cancer. The recruitment process was similar to that previously described for the case control study (2.2.5). The same questionnaire and blood samples were taken on the morning of surgery and processed as previously described (2.2.10).

2.3.3   Tissue Collection

Lung specimens were collected from surgical theatre immediately after resection and transferred to the Pathology Department on ice. A consultant pathologist was responsible for sampling the lung tissue. A random sample from normal lung tissue and tumour tissue were taken and placed in RNAlater RNA stabilisation reagent (Qiagen). Samples were then immediately stored at -80°C.

2.4   Laboratory Methods

2.4.1   Genomic DNA Extraction

2.4.1.1 Whole Blood and Lung Tissue

Genomic DNA was extracted from aliquots of clotted whole blood (1-2ml) using the Qiagen genomic DNA kit (Qiagen). Aliquots of blood were thawed at $37^{\circ}C$ for 15 minutes prior to centrifugation (2000g for 15 minutes) through a clotspin basket (Gentra). Baskets were washed with 7.5ml G2 lysis buffer to remove residual clot. G2 buffer (2ml) containing heat inactivated RNAase (100mg/ml) and $95\mu l$ of proteinase K was then added. The sample was vortex mixed and left rotating overnight at $4^{\circ}C$. The next day, the sample was incubated at $37^{\circ}C$ for 1 hour, vortex mixed and placed on ice. Qiagen genomic tips were prepared as per the manufacturers instructions. Buffer QBT (4ml) was run through each column prior to the sample being added. The column was then washed twice with 7.5ml G2 buffer. DNA was eluted using 5ml buffer QF, precipitated with 0.7 volume of isopropanol and centrifuged at 3500g for 15 minutes. The DNA pellet was transferred to an Eppendorf tube and washed twice in ice cold 70% ethanol, centrifuged at 3500g and the excess ethanol decanted to allow the DNA pellet to air dry. The pellet was resuspended in ddH_2O and the DNA concentration quantified (see section 2.4.2). Lung tissue was thawed and kept on ice at all times then homogenised in G2 lysis buffer. The lysate was then centrifuged at 15000g for 3 minutes and the supernatant was transferred to a Genomic tip, using the same protocol as whole blood.

2.4.1.2 Lymphocytes

Lymphocyte samples were rapidly defrosted and split into two aliquots. Two separate methods for DNA separation were used. DNA for measurement of N7-medG was separated from one lymphocyte aliquot using the same protocol as whole blood (Section 2.4.1.1). This method resulted in artefactual generation of oxidative damage. DNA for measurement of 8-oxodG was therefore separated from the second lymphocyte aliquot using the chaotropic method (Ravanat, Douki et al. 2002). This method produced an increase in non-specific signal in the immunoslotblot assay.

2.4.2 DNA Quantification

The concentration of DNA in samples was determined by fluorescence using a Picogreen dsDNA Quantitation Kit (Invitrogen). A standard curve was prepared using ultra pure calf thymus DNA (Sigma) in buffer I (concentrations 0, 0.028, 0.055, 0.11, 0.22, 0.44, 0.88 and 1.76µg/ml). A 10µl aliquot of each standard and unknown were added in duplicate to a black 96 well microtitre plate. Picogreen reagent (100µl) (10µl Picogreen dsDNA quantitation reagent, 9.5ml ddH$_2$O, 0.5ml 20xTE (10mM Tris, 1mM EDTA, pH 7.5)) was then added to each well. The plate was read immediately on a TECAN GENios plate-reader (plotted standard curve of fluorescence against concentration (µg/ml)). Results were accepted if the standard curve was linear (r > 0.99), the unknown values were between 0.1 and 1.5µg/ml and the coefficient of variance of the duplicates was less than 5%.

2.4.3 Protein Quantification

The concentration of protein in samples was determined by spectrophotometry using Bio-Rad protein assay reagent. A standard curve was prepared from BSA diluted in buffer I to concentrations of 0, 0.02, 0.03, 0.04, 0.05, 0.06, 0.08 and 0.1mg/ml. Bio-Rad reagent was diluted to 1x solution and filtered (2x 3MM filter paper) prior to use. A 40µl aliquot of each standard and unknown (diluted as appropriate in buffer I) were added in duplicate to a microtitre plate. Bio-Rad reagent (200µl) was added to each well and left at room temperature for 5 minutes. The plate was read at 595nm on a TECAN GENios plate reader. The results were accepted when the standard curve was linear (r > 0.99), the unknown values fell within the range of the standard curve (0.015 – 0.075mg/ml) and the coefficient of variance was less than 5% between duplicates.

2.4.4 RNA Extraction

2.4.4.1 Blood

Blood (2.5ml) for RNA extraction was collected into PAXgene Blood RNA tubes (PreAnalytiX) and stored at -80°C. Extraction was performed using the PAXgene Blood

RNA kit following instructions from the manufacturer; an optional on-column DNase digestion using the RNase-Free DNase Set was used (Qiagen). Blood was thawed and incubated at room temperature overnight to allow for cell lysis; then centrifuged (4000g) at room temperature for 10 minutes. The supernatant was removed and 4ml RNase-free water added, the sample was vortex mixed and again centrifuged (4000g) at room temperature for 10 minutes. The supernatant was removed and 350µl BR1 buffer added prior to vortex mixing. The sample was transferred into a microcentrifuge tube and 300µl buffer BR2 and 40µl proteinase K were added. After vortex mixing the sample was incubated for 10 minutes at 55°C, then vortex mixed for 30 seconds and centrifuged at 15000g for 10 minutes. The supernatant was carefully placed in a new microcentrifuge tube and 350µl ethanol was added. Two 700µl aliquots of sample were added into the PAXgene RNA spin column sequentially and the flow-through discarded after centrifugation (1 minute at 20 000g). The spin column was then washed with 350µl buffer BR3 and spun at 20000g for 1 minute. 80µl DNase 1 incubation mix (10µl DNase 1 stock and 70µl DNA digestion buffer) was added to the spin column and incubated at room temperature for 15 minutes. The column was then washed once with 350µl buffer BR3 and twice with 500µl of buffer BR4. The spin column was then placed in a microcentrifuge tube and 40µl elution buffer (BR5) was added before centrifugation for 1 minute at 20000g. The elution step was then repeated. The eluate was incubated at 65°C for 5 minutes, then placed on ice. RNA was quantified using a Nanodrop ND-100 Spectrophotometer (Labtech) and purity was assessed by running samples 500ng on a 1% agarose gel stained with ethidium bromide. Visualisation of sharp 28S and 18S bands was required, before the sample was accepted and stored at -80°C.

2.4.4.2 Lung and Tumour Tissue

Lung and tumour tissue for RNA extraction were placed immediately into RNAlater RNA stabilisation reagent (Qiagen) prior to storage at -80°C. When required, samples were removed from -80°C storage and thawed at room temperature. A section of tissue was removed with a scalpel and weighed. Extraction was performed using RNeasy Mini Kit (Qiagen). The tissue was disrupted and homogenised in buffer RLT (+ 10µl / ml β-mercaptoethanol) by passing the lysate through a 20 gauge needle attached to an RNase free syringe five times. The tissue lysate was then centrifuged at 15000g for 3 minutes and the supernatant transferred to an RNase free 2ml eppendorf. A 600µl aliquot of 70% ethanol was added immediately and 700µl of sample was then added to an RNeasy mini

column prior to centrifugation at 10000rpm for 15 seconds at room temperature. RW1 buffer (350μl) was added to the RNeasy mini column and centrifugation repeated. DNase 1 incubation mix (80μl) (10μl DNase stock solution + 70μl buffer RDD) was directly added to the RNeasy silica gel membrane and left at room temperature for 15 minutes. Buffer RW1 (350μl) was added to the column and centrifugation repeated. Buffer RPE (500μl) was then added and the column was then centrifuged. A second 500μl aliquot of buffer RPE was added to the tube and centrifuged for 2 minutes at 10000 rpm to dry the membrane. To elute the tube 30μl of RNase free water was added to the column and the flow through was collected during centrifugation at 10000rpm for 1 minute. A 500ng aliquot of each sample was run on an agarose gel to check for DNA contamination, prior to storage at -80°C.

2.4.5 Genotyping

The genotyping of hOGG1 and MGMT was performed by G McGown and M Thorncroft. All genotyping was performed blinded to the case-control status of the samples.

2.4.5.1 Site Specific PCR Amplification of DNA

The specific gene fragments to be studied were amplified using PCR techniques (Tables 2.2 and 2.3). A master mix was made from the following materials: dNTPs (0.25mM of each), 10x Taq buffer, ddH$_2$O, forward and reverse primers (0.1mM of each) (Table 2.2). Sample DNA (100ng) was then added and layered with two drops of mineral oil. A standard hot start was used; this included an initial denaturing step (95°C for 5 minutes) followed by a holding step at 80°C, during which time the *Taq* DNA polymerase was added as part of a hot start mix (consisting of Taq 0.2U, 1x Taq buffer and H$_2$O). Amplification of DNA was then achieved using thermal cycles detailed in table 2.3. Primary PCR products were resolved using 2% agarose gels stained with ethidium bromide and visualised using Genesnap (Syngene) software.

2.4.5.2 GSTM1 and GSTT1

The determination of the GSTM1 and GSTT1 genotype and the presence of the CYP2D6 fragment, which was used as an internal control, was undertaken using methods previously described in this population (Lewis, Cherry et al. 2002).

Table 2.2 Sequence of oligonucleotide primers used for PCR amplification of selected genotypes.

Gene	Sequence of Oligonucleotide Primers		Product Size (bp)
	Forward (5'→3')	Reverse (5'→3')	
GSTM1	CTGCCCTACTTGATTGATGGG	CTGGATTGTGAGCAGATCATGC	273
GSTT1	TTCCTTACTGGTCCTCACATCTC	TCACCGGATCATGGCCAGCA	480
CYP2D6	GCCTTCGCCAACCACTCCG	AAATCCTGCTCTTCCGAGGC	334
hOGG1 codon 326	ACTGTCACTAGTCTCACCAG	CCTTCCGGCCCTTTGGAAC	156
MGMT rs12268840	CACCAAGTACCATCAGCCAC	GAGAGTACCATAGTCTAAGGGGC	254
MGMT codon 178	TCCATGCTGAGACATAGCTGAC	TGAGCTCCCTCCCAAGC<u>A</u>GG	307

Table 2.3 PCR conditions for selected genotypes.

Gene	PCR Conditions						Cycles
	Denaturing		Annealing		Extending		
GSTM1	45s	94°C	45s	62°C	60s	72°C	35
GSTT1							
CYP2D6							
hOGG1	60s	94°C	60s	60°C	60s	72°C	35
MGMT rs12268840	60s	94°C	60s	63°C	60s	72°C	35
MGMT codon 178	60s	94°C	60s	66°C	60s	72°C	35

A blank and two positive controls (samples b6 and t10) were used for each PCR amplification run (Figure 2.1). The positive control samples were taken from a previous study (unpublished data) and had therefore been identified as GSTM1 –ve / GSTT1 +ve for sample b6 and GSTM1 +ve / GSTT1 +ve for sample t10. Samples were run on 2% agarose gel (100ml 1x TBE, 2g agarose) stained with ethidium bromide (genotypes shown are: b6 = GSTM1-/T1+; t10 = M1+/T1+; 37, 41 and 47 = M1+/T1+; 39, 42 and 44 = M1-/T1+; 45 = M1+/T1-).

blank b6 t10 37   39   41   42   44   45   47   sample identification

GSTT1 (480bp)

CYP2D6 (334bp)

GSTM1 (273bp)
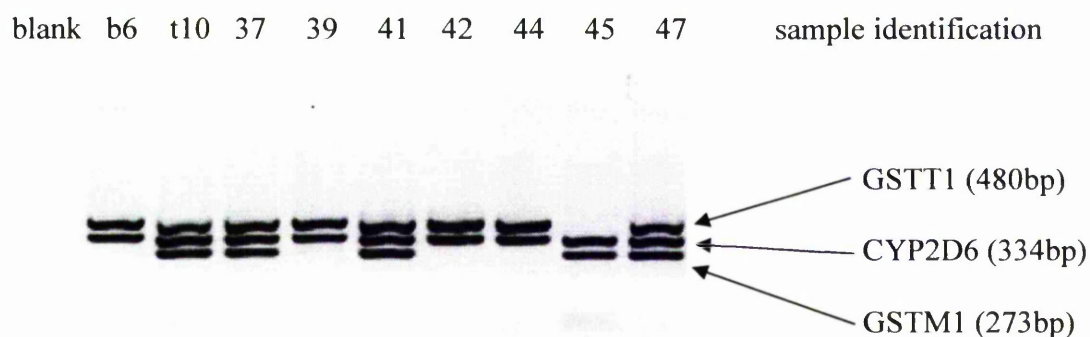
Figure 2.1 Visualisation of GSTM1, GSTT1 and CYP2D6 primary PCR products after separation using agarose gel electrophoresis.

2.4.5.3 MGMT rs12268840 (Intron 1)

Digestion of the primary PCR product was performed using the restriction enzyme *BsaJ*1 (New England Biolabs). 2.5µl of 10x NEBuffer 2, 2.1µl H_2O and 0.4µl of *BsaJ*1 (10 U/µl) were added to 20µl of PCR product and incubated for 2hr at 60°C. The enzyme was heat inactivated at 80°C for 20 min. 20µl of digestion reaction with 5µl 5x orange loading buffer (Trevigan) was run on a 3% MetaPhor gel (Cambrex) containing 0.1% ethidium bromide. There were three restriction sites within the 254bp PCR product, two of which were polymorphic; when no polymorphisms were present four digestion products were created (Figure 2.2). Genotyping of the rs12268840 polymorphism was determined by analysing the combination of digestion products detailed in table 2.4 (Margison, Heighway et al. 2005). The 254bp intron 1 (rs12268840) PCR product contained 3 restriction sites for the enzyme *BsaJ*1. Two of the sites were polymorphic, multiple permutations of digestion products were therefore produced. Limiting the analysis to the presence or absence of the 41, 77 or 140bp products (Table 2.4) enabled a more simplistic determination of the rs12268840 genotype (114bp product was present in all samples) (Figure 2.3).
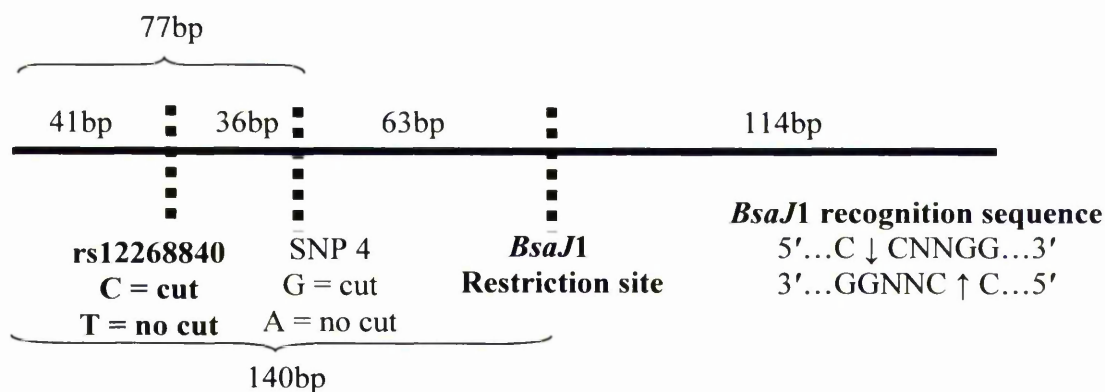
Figure 2.2 *BsaJ*1 recognition sequence and restriction sites within the rs12268840 (intron 1) primary PCR product.

Table 2.4 Initial screen used to determine the rs12268840 genotype.

rs12268840 Genotype	Presence (+) or Absence (-) of Digestion Products		
	41bp	77bp	140bp
CC	+	-	-
CT	+	+	-
	+	-	+
TT	-	+	-
	-	-	+
	-	+	+

Figure 2.3 Visualisation of rs12268840 primary PCR products after *BsaJ*1 restriction digestion.

78

2.4.5.4 MGMT Codon 178

Restriction digestion of the primary PCR product was performed using the restriction enzyme *Stu*1 (Promega). 2.5µl of 10x Buffer B and 1µl of *Stu*1 (10 U/µl) were added to 21.5µl of PCR product and left for 2hr at 37°C. The enzyme was heat inactivated at 65°C for 20 min. 20µl of digestion reaction with 5µl 5x orange loading buffer (Trevigan) was run on a 3% MetaPhor gel containing 0.1% ethidium bromide. The reverse primer introduced a mismatched base (A to T) and so created a restriction site for the enzyme *Stu*1 in the presence of a G residue (R allele); resulting in a 286bp and a 21bp digestion product (Figure 2.4). No enzyme restriction site was present with the A residue resulting in an uncut 307bp product (K allele) (Margison, Heighway et al. 2005).

(a)

*Stu*1 recognition sequence	5'...AGG ↓ CCT...3'
	3'...TCC ↑ GGA...5'

(1)   3' – GG<u>A</u> CCG AAC CCT CCC TCG AGT – 5'

(2) 5' – CAC CGG TTG GGG **A(A/G)G** CCA GGC TTG GGA GGG AGC TCA – 3'

Codon 178

(b)
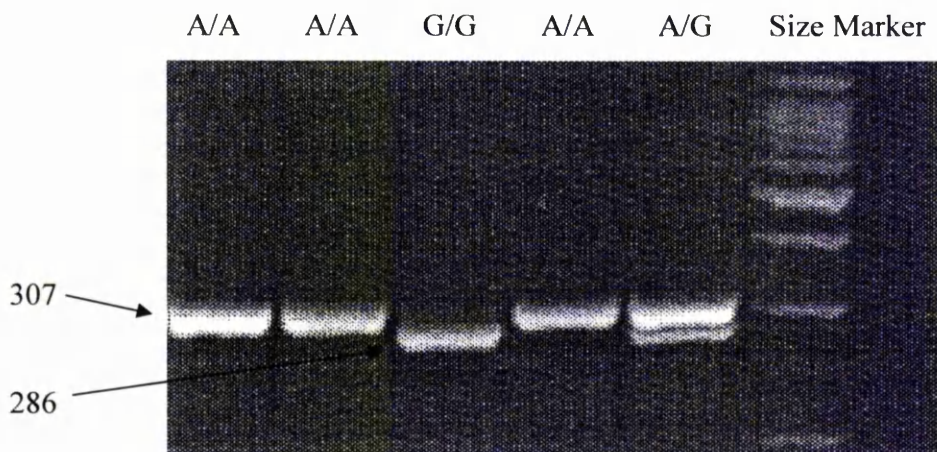
Figure 2.4 (a) *Stu*1 recognition sequence and the introduction of a mismatched base (A) by the reverse primer (1), introducing a restriction site in the presence of a G residue in codon 178 (2) – representing the R allele. (b) Visualisation of codon 178 primary PCR products after Stu1 restriction digestion.

2.4.5.5 hOGG1 Codon 326

hOGG1 genotype was determined using the method described by Chen *et al* (Chen, Elahi et al. 2003). Restriction digestion of the primary PCR product was performed at $37^{\circ}C$ for 2 hours in a 25µl reaction consisting of: 20µl primary PCR product, 0.1µl *Ita*1 (Roche, 10U/µl), 2.5µl buffer H and 2.4µl ddH$_2$O. The reaction was inactivated by heating to $65^{\circ}C$ and the digestion products run on a 3% MetaSeive agarose gel (Flowgen) stained with ethidium bromide for 1 hour.

No restriction digestion took place in the presence of a hOGG1 codon 326 Ser/Ser homozygous genotype; this resulted in the presence of a single 156bp fragment. Restriction digestion of Ser/Cys heterozygotes produced fragments sized 156, 100 and 56bp. In the presence of Cys/Cys homozygotes 100 and 56bp fragments were produced (Figure 2.5).

(a)

*Ita*1 recognition sequence   5'...GC ↓ NGC...3'

                                     3'...CGN ↑ CG...5'

(b)
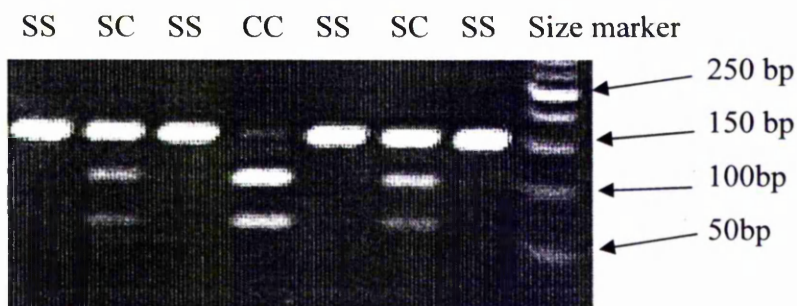
Figure 2.5 (a) *Ita*1 recognition sequence and (b) visualisation of hOGG1 codon 326 primary PCR products after *Ita*1 restriction digestion.

2.4.6 Oligonucleotide Cleavage Assay

2.4.6.1 Oligonucleotide Labelling and Annealing

All oligonucleotides used in the cleavage assay were synthesised by MWG and diluted to 1 pm/μL in ddH2O; the sequences used are listed in table 2.5. Substrate oligonucleotides containing enzyme specific DNA adducts were 5′ end labelled with [γ-^{32}P]-ATP (10μCi/mM, Amersham Biosciences). The labelling reaction (10μl) consisted of 2.5μl of oligonucleotide (1pmole/μl), 1μl of PNK buffer, 1.5μl of ddH2O, 4μl of [γ-^{32}P]-ATP (6.7 pmol) and 1μl of PNK. The labelling reaction was incubated at 37°C for 30 minutes. A complementary strand, which was 5′ end labelled with biotin was annealed to the substrate oligonucleotide by the addition to the labelling reaction of a mixture of 7.5μl oligonucleotide complement (1pmole/μl), 2.5μl 0.5M NaCl and 5μl ddH2O. The annealing reaction was incubated for 5 min at 80°C then allowed to cool to room temperature.

Table 2.5 The sequence and position of adducts in the cleavage assay oligonucleotides.

Oligonucleotide Name		Oligonucleotide Sequence (5′ to 3′) (X = position of adduct)
MPG	Substrate	CCG CTX GCG GGT ACC GAG CTC GAA T **X = ethenoadenine**
	Complement	biotin A TTC GAG CTC GGT ACC CGC TAG CGG
MGMT	Substrate	GAA CTX CAG CTC CGT GCT GGC CC **X = O^6-methylguanine**
	Complement	biotin GGG CCA GCA CGG AGC TGC AGT TC

2.4.6.2 Microtitre (Reaction) Plate Preparation

Streptavidin coated microtitre plates (96 well) (Thermo) were blocked for at least two hours at room temperature with 400μl/well of 0.5% BSA/PBS. Blocking solution was then aspirated and 12 fmole of oligonucleotide were added to each well (100μl of 25μl annealed oligonucleotide diluted in 21ml 0.5% BSA/PBS). Plates were stored overnight at 4°C. Bound oligonucleotide was measured in four wells using a scintillation counter the next

day and the binding efficiency calculated in reference to the number of counts expected (1209 Rackbeta liquid scintillation counter, LKB Wallace, counting efficiency 59%).

2.4.6.3 Sample Preparation and Sonication

All samples were transferred from $-80^{\circ}C$ storage and kept on ice at all times. PBMCs were rapidly defrosted at $37^{\circ}C$. The nuclear PBMC samples were centrifuged at 19800g for 10 minutes at $4^{\circ}C$, the supernatant decanted and 1ml of buffer I + leupeptin (10mg/ml) added to the pellet. Cellular PBMC samples were prepared by adding a 0.1 volume of 10x buffer I (+ leupeptin) to 0.9 volume of PBS containing the cellular suspension. Lung samples were sliced to an appropriate size, weighed, placed in 1ml of buffer I + leupeptin and manually homogenised with an eppendorf micropestle. An MCF-7 cell pellet (~5 million cells / aliquot) was prepared as a control for each assay by the addition of 1ml buffer I + leupeptin. A Heat Systems XL2020 ultrasonic processor with 3.2mm microtip was used for sonication. Prior to each set of samples, the temperature increase of 3 x 1ml ice cold aliquots of IBSA were measured with a thermocouple thermometer pre and post sonication. The accepted temperature change was $25 \pm 3^{\circ}C$, if the temperature was outside this range sonication time was increased or decreased to recalibrate to the $25 \pm 3^{\circ}C$ range. Samples were sonicated for 10s followed immediately by the addition of 10μl of PMSF in ethanol (final concentration 87μg/ml). The sonicate was then centrifuged at 19800g at $4^{\circ}C$ for 10min. The supernatant was transferred to a fresh Eppendorf tube on ice. The DNA and protein concentrations (see section 2.4.2 and 2.4.3) of the sample were then determined. Aliquots of the samples were diluted in IBSA to a concentration of 10μg DNA/ml.

2.4.6.4 MGMT Activity Assay

Samples were serially diluted (1:1) in IBSA in a microtitre plate to produce the following DNA concentrations: 10.00, 5.00, 2.50, 1.25, 0.63, 0.31 and 0.16μg/ml. The final well, which contained only IBSA and no extract, was the negative control. The reaction plate was washed five times with PBS and 50μl aliquots of IBSA added to each well. A 50μl aliquot of each DNA concentration from each sample was added to the reaction plate. The plate was incubated for 4 hours at $37^{\circ}C$. After the incubation period the plate was washed

again five times in PBS and the restriction enzyme *Pst*1 (100μl containing: 88μl ddH$_2$O, 1μl BSA 10x, 1μl *Pst*1, 10μl buffer 3) added to each well. The plate was incubated for one hour at 37°C before the supernatant was transferred to a LumaPlate (Packard), which was left to dry overnight. The LumaPlate was then read in a TopCount microplate scintillation counter (Packard).

2.4.6.5 MPG Activity Assay

Samples were serially diluted (1:1.3) to produce the following concentrations of DNA: 10, 7.7, 5.9, 4.6, 3.5, 2.7 and 2.1μg/ml. The final well, which contained only IBSA and no extract, was the negative control. The reaction plate was washed five times with PBS and 90μl aliquots of magnesium cleavage buffer added to each well. A 10μl aliquot of each DNA concentration from each sample was added to the reaction plate. The plate was incubated for 2 hours at 37°C. After incubation the supernatant was transferred to a LumaPlate, allowed to dry overnight and then read in a TopCount microplate scintillation counter.

2.4.6.6 Calculation of DNA Repair Activity

DNA repair activity was calculated for each sample by plotting, for all dilutions and the negative control, amount of DNA (μg) in an individual well against the measured disintegrations / minute (dpm) released from that well. The gradient of the line was calculated by linear regression and was only accepted if the R^2 value was ≥ 0.985. The gradient represented dpm of adduct cleaved / μg of DNA. On the reference date of the isotope, 13 200 dpm was equivalent to 1 fmole of ATP. This value was corrected according to the difference in time between the samples being measured in the TopCount microplate scintillation counter and the reference date of the isotope. The corrected dpm / fmole value was then divided into the dpm measured from the wells to produce a value with units fmole (of adduct cleaved) / μg DNA. For the MPG assay this value was then divided by 2 to produce a value / hour. The final units for the MPG assay were fmole / μg DNA / hr. The MGMT protein is inactivated by the removal of one adduct therefore the final units were simply fmole (adduct cleaved or active MGMT molecules) / μg DNA.

2.4.7 Polyacrylamide Gel Electrophoresis

Polyacrylamide gel analysis was performed using MiniProtean (BioRad) apparatus. A 20% gel (10ml/gel) was prepared using 5ml Accugel (40%, Mensura), 1ml 10x TBE (500mM Tris-borate, 10mM EDTA, pH 8.0), 4ml ddH$_2$O followed by the addition of 100μL 10% ammonium persulphate and 10μL TEMED. Following gel polymerisation the wells were washed and then placed in TBE. Samples were loaded with a non-denaturing loading buffer (30% glycerol, 0.25% xylene cyanol, 0.25% bromophenol blue) and run at 200V observing the separation of xylene cyanol (45bp) and bromophenol blue (12bp). After approximately 45 minutes the gel was removed, wrapped in Saran and placed against a phosphorimaging screen which was later read in a STORM phosphorimager (BioRad).

2.4.8 N7-Methyldeoxyguanosine Determination

N7-MedG levels were determined, by K Harrison, from DNA extracted from both lymphocytes and lung tissue. The immunoslotblot method, as described by Harrison *et al*, was used (Harrison, Wood et al. 2001). In brief, sonicated DNA from samples (in triplicate) and MNU-methylated calf thymus DNA standards were alkali treated to open the imidazole ring of N7-medG adducts in the DNA (70mM NaOH, at 37°C for 30 minutes). Samples and standards were then neutralised, heat denatured (5 minutes at 100°C) and cooled on ice for 10 minutes to generate single stranded DNA. Samples were then immobilised on nitrocellulose filters and the slots rinsed. Filters were removed from the support and baked at 80°C for 90 minutes and then blocked for 1 hour with PBS-Tween (0.1% v/v) containing 5% fat free milk powder. Primary rabbit polyclonal antibody diluted 1:10000 in PBS-Tween (0.1% v/v) containing 0.5% fat-free milk powder was applied overnight at 4°C incubation. The nitrocellulose filters were washed and then incubated with goat anti-rabbit IgG horseradish peroxidase conjugate diluted 1:10000 in PBS-Tween (0.1% v/v) containing 0.5% fat free milk powder for 1 hour at room temperature.

The enzymatic activity associated with slots was visualised by bathing the nitrocellulose filters in chemiluminescent reagent and exposing to X-ray film. The band intensities were quantified by scanning the X-ray films with an optical scanner (Storm 860; Amersham Biosciences). The assay limit of quantitation was 0.1 N7-medG / 10^{6}dG.

2.4.9 8-Hydroxydeoxyguanosine Determination

8-oxodG levels in DNA extracted from lymphocytes was determined by HPLC-ECD, by M Carus; materials and methods described by Harrison *et al* (Harrison, Crosbie et al. 2006). In brief, up to 50µg of DNA was hydrolysed in 110µl (final concentrations: 4U Nuclease P_1, 1U alkaline phosphatase, 25mM sodium acetate, 0.05mM zinc chloride, pH 5.3) for 1 hour at 50°C. Following centrifugation (13000g for 5 minutes) the upper phase was transferred to a recovery vial and injected by a Waters 717 Plus autosampler (Waters) onto a Supercosil LC-18-DB column (150x4.6mm, 3µm particle, Superco). The HPLC conditions were modified from a method by Hofer *et al* (Hofer and Moller 2002). These consisted of 3% aequeous methanol containing polished water (C18 SPE column), 50mM (mono basic) potassium filtered buffer (pH 5.5) at an isocratic flow of 1 ml/minute. UV detection (Perkin Elmer LC90 UV) was at 254nm and data integrated by a Shimadzu C-R113 integrator and an electrochemical detector ((EC) CoulArray, ESA Analytical Ltd) used a 4-cell system, with the cells operated at 200, 260, 320 and 380mV. The limit of quantitation of 8-oxodG in human DNA was ~ 50fmol (3 times background). Due to ECD response variation, calibration standards of 8-oxodG were run with each analysis as were the calibration standards of dG (CV 3.8%). The correlation coefficient was 0.98-1.00 for each calibration curve. A calf thymus DNA was used as a standard for assay comparison: intrabatch and interbatch reproducibility was 2% and 11% respectively. 8-oxodG was dissolved in HPLC water and the concentration checked by measuring the absorbance at 248nm and dividing by the extinction coefficient. After dilution 100µl aliquots (10nM) were frozen as stocks, these were then used to generate standards (50 – 1000 fmol / injection volume). Raji cells were obtained from Cancer Research UK fermentation services, aliquoted and frozen (-80°C) as cellular pellets containing approximately 10^7 cells until use.

2.4.10 Determination of MGMT Relative Allelic Expression

The determination of allelic expression imbalance was performed by G MCGown and M Thorncroft.

2.4.10.1 Reverse Transcription of RNA to Produce cDNA

cDNA was created by reverse transcription PCR (RT-PCR) of RNA using a Reverse Transcription System (Promega) and following the manufacturers protocol. A 1µg aliquot of sample RNA was placed in a microcentrifuge and incubated at 70°C for 10 minutes. After brief centrifugation the sample was placed on ice. A 20µl reaction mix was made containing (final concentrations) 5mM MgCl$_2$, 1x reverse transcription buffer, 1mM each dNTP, 1U/µl Recombinant RNasin® Ribonuclease Inhibitor, 15U/µg AMV Reverse Transcriptase (High Conc.), 0.5µg Oligo(dT)$_{15}$ primer and 1µg sample RNA. The reaction was incubated at 42°C for 15 minutes, then heated at 95°C for 5 minutes and then incubated at 5°C for 5 minutes.

2.4.10.2 Site Specific PCR Amplification of Exon 5

2.4.10.2.1 cDNA

The primary PCR amplified 100ng of cDNA. The master mix contained (final concentrations) 0.5µM each primer (Exon 3 forward CTGCTGCACAGCTAGTTGAGAC and exon 5 reverse TGAGCTCCCTCCCAAGCCAGG), 1x Taq buffer (Promega), 100µM dNTPs and 1.5mM MgCl$_2$. A standard hot start was used by heating samples to 95°C for 5 minutes then an 80°C hold, before the addition of 2.5U Taq (Promega). There were then 35 cycles consisting of one minute each of 94°C, 66°C and 72°C. This was followed by 10 minutes at 72°C and finally a 4°C hold. A secondary PCR was performed using 1µl of diluted (1 in 100) primary PCR product in a 50µl reaction. The conditions were the same as for the primary run except 1U of Taq was added and there were 10 cycles. The PCR product was 367bp, samples were run on 2% agarose gels to check the size. Restriction digestion was performed by *Stu*1 (Promega), the digestion reaction consisted of 15µl PCR product, 2µl 10x Buffer 2 (Promega), 1U *Stu*1 and 2.9µl ddH$_2$O. The reaction was incubated at 37°C for 2.5 hours and then heat inactivated at 65°C for 20 minutes. Digestion

of exon 5 heterozygotes produced the following bands: 367, 345 and 22bp. A control was used for each digestion run to confirm complete digestion (sample 446 – exon 5 homozygote).

2.4.10.2.2 Genomic DNA

The PCR mixes and conditions for both primary and secondary PCR were the same as those described for cDNA (Section 2.4.10.2.1). The primers used were exon 5 forward TCCATGCTGAGACATAGCTGAC and exon 5 reverse TGAGCTCCCTCCCAAGCCAGG. The PCR product was 367bp, samples were run on 2% agarose gels to check the size. Restriction digestion was performed by *Stu*1 (Promega) using the conditions as previously described (Section 2.4.10.2.1). Digestion of exon 5 heterozygotes produced 367, 345 and 22bp bands. A control was used for each digestion run to confirm complete digestion (sample 446 – exon 5 homozygote). The uncut product related to the K allele and the cut 345 product to the R allele.

2.4.10.3 Agilent Analysis of Allelic Imbalance

Digestion products (1μl) were run on DNA 500 chips in an Agilent 2100 Bioanalyzer (Agilent Technologies). 2100 Expert Software (Agilent Technologies) was used to analyse the results. Relative imbalance was calculated by determining the area under the peak at the expected size of product (Figure 2.6).

(a)   Agilent pseudogel electrophoretogram of MGMT (codon 178) mRNA PCR products after Stu1 digestion.

(B = whole blood, L = normal lung tissue, T = tumour tissue)

Ladder s11 s23 s11 s11 s19 s19 s23 s23 s10 446 s19 empty
(600bp) B   B   L   T   L   T   L   T   T   B   B
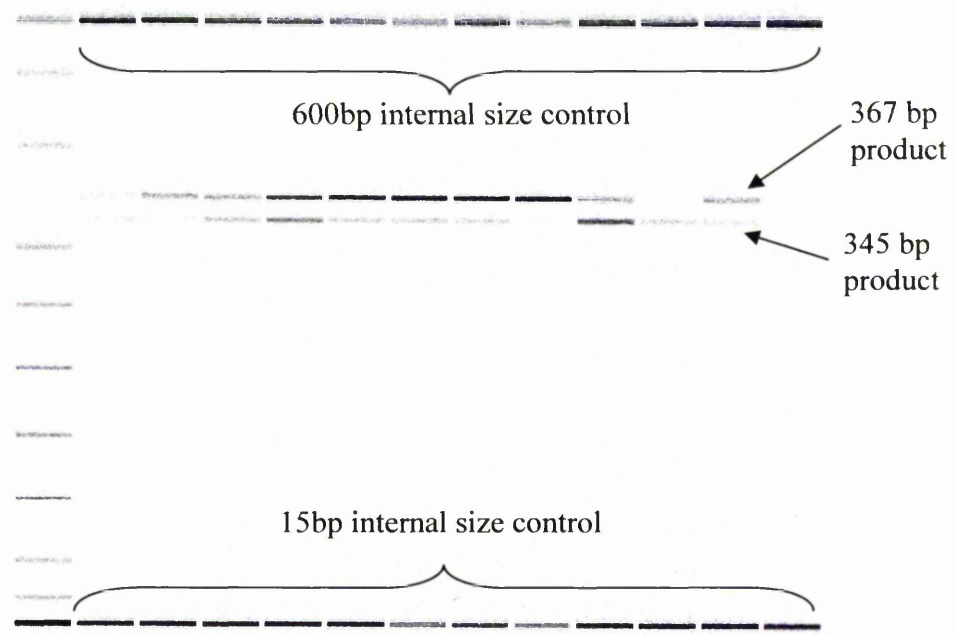
(b)   Agilent trace of relative allelic expression (s23 – whole blood ratio 3.1K:1R).
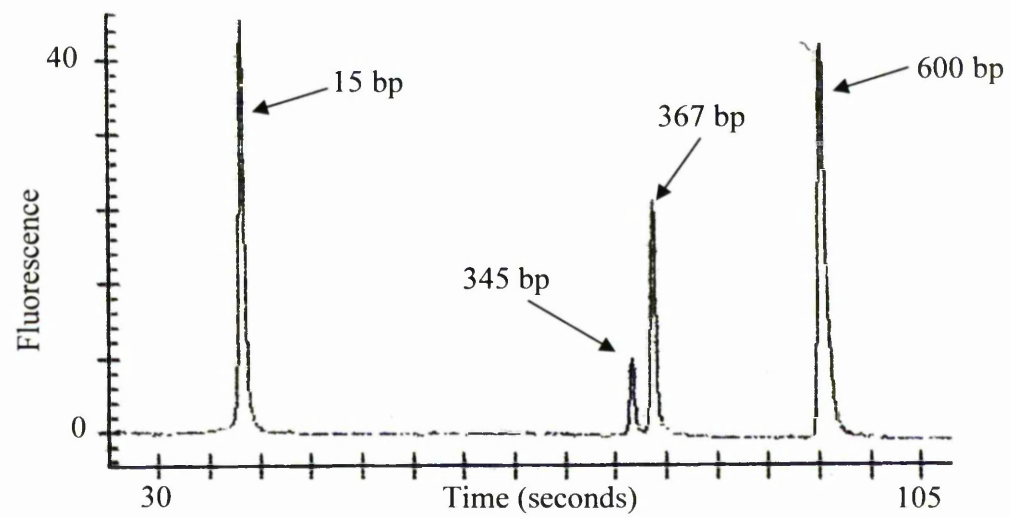
Figure 2.6   Examples of (a) an agilent pseudogel electrophoretogram and (b) an agilent trace, used in the determination of relative allelic expression of MGMT. (codon 178 – 367bp product = K178 allele, 345bp product = 178R allele)

2.4.11 Real Time PCR Determination of GSTM1 Gene Copy

GSTM1 gene copy number was determined in all subjects who were designated GSTM1 positive after initial genotyping as described in section 2.4.5.2. A multiplex real time PCR was performed using an ABI Prism® 7900HT Sequence Detection System (Applied Biosystems) in a 384 well format. The method used was based on that described by Brasch-Anderson *et al* (Brasch-Andersen, Christiansen et al. 2004). The primers and probes used are detailed in table 2.6, the authors state that the amplification efficiencies of the genes were approximately equal. Primers and probes were obtained from MWG.

Table 2.6 GSTM1 and albumin primers and probes used for real time PCR.

		Sequence (5′-3′)
GSTM1	forward primer	TGGACATTTTGGAGAACCAGACC
	reverse primer	TCACAAATTCTGGATTGTAGCAGAT
	probe	**Fam**-ATGGACAACCATATGCAGCTGGGCAT-**Tamra**
Albumin	forward primer	GCTGTCATCTCTTGTGGGCTGT
	reverse primer	ACTCATGGGAGCTGCTGGTTC
	probe	**Joe**-CCTGTCATGCCCACACAAATCTCTCCC-**Tamra**

All assays were performed on 384 well optical reaction plates (Applied Biosystems) with a 10µl PCR volume, using a epMotion 5070 (Eppendorf) robot to aliquot. The PCR reaction contained 5µl of 2x TaqMan® Universal PCR Master Mix, 150nM GSTM1 primers and probes, 200nM albumin primers and probes and 40ng of sample genomic DNA. Each sample and each assay was performed in triplicate. PCR conditions were as follows: initial 2 minute hold at 50°C then 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Results were analysed using Sequence Detection Software (v. 2.1).

The relative number of gene copies was determined by the delta (Δ) C_T method. The threshold cycle (C_T) was defined as 10 standard deviations above the mean of baseline fluorescence which was calculated between cycles 3 to 15. The C_T of GSTM1 was subtracted from the C_T of albumin to produce the ΔC_T value. A mean ΔC_T value was calculated for each sample from three independently calculated ΔC_T values.

Chapter Three

Case-Control Study
Population Characteristics

3.1   Case-Control Study Recruitment

Study recruitment from the Bronchoscopy Unit at Wythenshawe Hospital took place between March 2003 and June 2004. During this time 897 patients underwent bronchoscopy, 425 of whom were eligible for study participation based on the inclusion and exclusion criteria described in section 2.2.3. Of those asked, 92.7% consented to participate in the study (n = 394). The remaining 472 patients were made up of those who were ineligible to participate (n = 337) and those who were not approached because the investigator was not present at the bronchoscopy list (n = 135). The reasons for patient exclusion are detailed in Figure 3.1; the most common reason was a previous diagnosis of any cancer (n = 167). Ten patients, who had consented to study participation, were excluded after enrolment because of the subsequent diagnosis of either a malignant or benign tumour that was not lung cancer. These diagnoses were: mesothelioma (3), melanoma (2), hamartoma (2), laryngeal carcinoma (1), gastric cancer (1) and lymphoma (1). A second control group was recruited from patients, who met the inclusion and exclusion criteria (section 2.2.3), attending the Chest Clinic at Wythenshawe Hospital between July 2004 and December 2005. The specific clinic sampled was determined by the availability of the author. A total of 90 subjects, referred to as chest clinic controls, consented to study participation, this represented 95% of those individuals who were approached. A blood sample was taken from 99.2% (n = 381 / 384) of subjects recruited from the Bronchoscopy Unit and 100% recruited from the Chest Clinic. The three subjects, who were all cancer free, where blood was not taken, were not included in further analyses of the study population because no genotyping or DNA repair assays could be performed. The final Bronchoscopy Unit study population was made up of 381 patients, consisting of 164 cases and 217 cancer free controls, referred to as bronchoscopy controls (Table 3.1).

Table 3.1 The case-control study population.

Study Population	n
Cases	164
Bronchoscopy Controls	217
Chest Clinic Controls	90
TOTAL	**471**
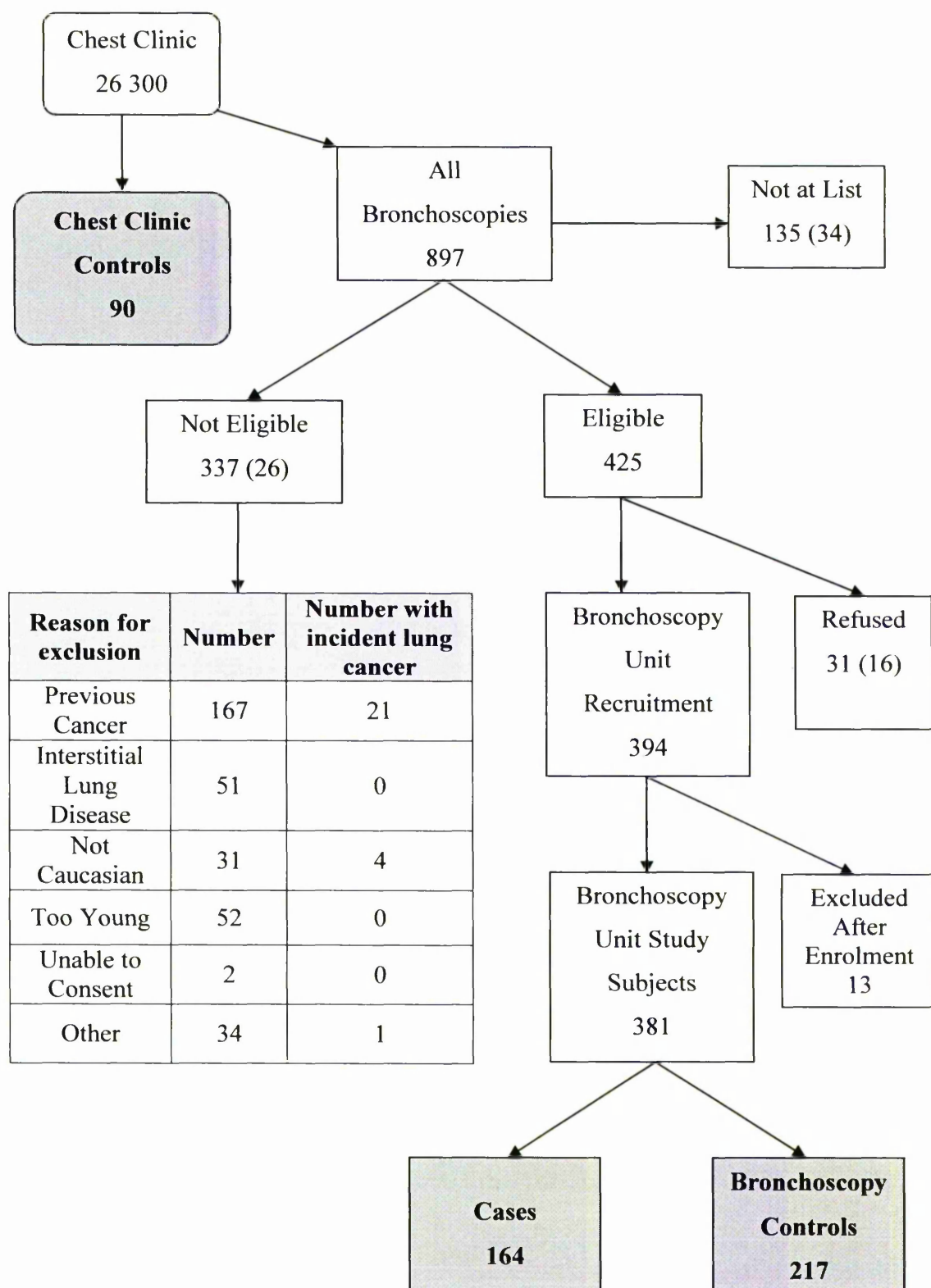
Figure 3.1. A flow chart of case-control study recruitment.

(Numbers in brackets represent patients diagnosed with incident lung cancer).

3.2   Diagnoses of Cases and Controls

3.2.1   Lung Cancer

A total of 164 subjects recruited to the study received a new diagnosis of lung cancer; these individuals were categorised as cases. This represented 68.3% of all incident diagnoses of lung cancer (n = 164 / 240) and 91.1% (n = 164 / 180) of eligible patients with lung cancer passing through the Bronchoscopy Unit during the recruitment period.

3.2.1.1 Histology and Stage

A histological diagnosis was achieved in 80.5% of cases (n = 132) with 19.5% (n = 32) diagnosed clinically. Cases without histology were significantly older than those with a histological diagnosis (74.1 ± 9.9 vs. 68.1 ± 10.1 years, p = 0.003). This discrepancy in age may be due to additional invasive diagnostic techniques e.g. surgical biopsy or repeat bronchoscopy being offered to a younger and possibly fitter age group. Overall, squamous cell carcinoma was the most common histological type of lung cancer, accounting for just over a quarter of cases (26.2%). This was followed by adenocarcinoma (21.3%), small cell carcinoma (18.9%) and unspecified NSCLC (12.2%) (Table 3.2).

Table 3.2 Lung cancer histology.

Histology of lung cancer	All Cases	
	n	%
Squamous Cell Carcinoma	43	26.2
Adenocarcinoma	35	21.3
Small Cell Carcinoma	31	18.9
Unspecified NSCLC	20	12.2
Tumour	2	1.3
Giant Cell Lung Carcinoma	1	0.6
None	32	19.5
Total	164	100

The distribution of squamous cell and adenocarcinoma differed according to gender. Males were twice as likely to be diagnosed with squamous cell carcinoma and half as likely to be diagnosed with adenocarcinoma as females (Table 3.3). There were no significant differences in distribution of small cell or unspecified NSCLC by gender. The probability of a histological diagnosis was the same in males and females (p = 0.57).

Table 3.3 Distribution of histological cell type stratified by gender.

Histology	Males n (%)	Females n (%)	OR[a] (95% CI)	p value
Squamous Cell Carcinoma	33 (29.7)	10 (18.9)	1.82 (0.82–4.05)	0.14
Adenocarcinoma	19 (17.1)	16 (30.2)	0.48 (0.22–1.03)	0.06
Small Cell Carcinoma	22 (19.8)	9 (17.0)	1.21 (0.51–2.84)	0.66
Unspecified NSCLC	13 (11.7)	7 (13.2)	0.72 (0.26–1.99)	0.53
Tumour	1 (0.9)	1 (1.9)	-	-
Giant Cell Lung Carcinoma	0	1 (1.9)	-	-
None	23 (20.7)	9 (17.0)	0.78 (0.33–1.83)	0.57
Total	111	53	-	-

[a]Odds of each histological cell type occurring in males (reference group = females).

To simplify analysis of the stage of lung cancer at diagnosis, the TNM staging classification was reduced to three categories. For NSCLC, stages Tx to 1B have been referred to as early stage, 2A to 3A as intermediate stage and 3B to 4 as advanced disease. Limited small cell carcinoma was categorised as early stage and extensive small cell carcinoma as advanced. Two thirds (67.1%) of cases were diagnosed with advanced disease (Figure 3.2 and Table 3.4).

☐ Early
▨ Intermediate
■ Advanced
☐ Unknown
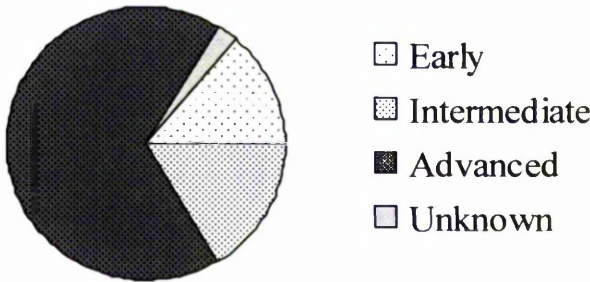
Figure 3.2 The frequency of lung cancer stage at time of diagnosis.

Table 3.4 Lung cancer stage at diagnosis.

NSCLC		n	%	Small Cell		n	%
Stage		n	%	Stage		n	%
Early	(Tx – 1B)	22	16.5	Early	(Limited)	1	3.2
Intermediate	(2A – 3A)	27	20.3				
Advanced	(3B – 4)	80	60.2	Advanced	(Extensive)	30	96.8
Unknown		4	3.0				
Total		133	100	Total		31	100

Cases with no histology were diagnosed in the earlier stages of lung cancer more frequently than histologically confirmed cases (p = 0.01) (Table 3.5). Subjects diagnosed with small cell carcinoma almost invariably presented with advanced disease (96.8%), only one case had early stage disease (Table 3.4 + 3.5). A high proportion of cases with adenocarcinoma (82.4%) also presented late. Squamous cell carcinoma had a lower proportion of patients presenting with advanced disease (53.7%) and an increased proportion presenting with intermediate stage disease (36.6%). This may reflect the fact that squamous cell carcinoma tends to grow in the proximal lung and therefore cause symptoms such as haemoptysis earlier, leading to earlier diagnosis.

Table 3.5 The distribution of lung cancer stage in cases according to histological type.

Cases		Stage of Lung Cancer at Diagnosis						p value
		Early		Intermediate		Advanced		p value
		n	%	n	%	n	%	
All with histology		14	10.9	20	15.5	95	73.6	0.01
No Histology		9	29.0	7	22.6	15	48.4	
Histological Type	Small Cell	1	3.2	0	0	30	96.8	
	Adenocarcinoma	5	14.7	1	2.9	28	82.4	
	Squamous Cell	4	9.8	15	36.6	22	53.7	
	Large Cell	4	22.2	2	11.1	12	66.7	
	Other	0	0	2	40.0	3	60.0	

3.2.2 Benign Respiratory Disease

A positive history of non-malignant respiratory disease, which included prevalent diagnoses and excluded previous resolved conditions, was reported in 38.4% of cases (n = 63), 51.6% of bronchoscopy controls (n = 112) and 87.8% of chest clinic controls (n = 79). A history of chronic respiratory disease was therefore significantly more common in both bronchoscopy (p = 0.01) and chest clinic controls (p = 0.0003) (Table 3.6). Chronic obstructive pulmonary disease (COPD) or emphysema was present in approximately one quarter of cases and a similar frequency of bronchoscopy controls but was significantly more common in chest clinic controls (p = 0.0009). Asthma was a less frequent diagnosis but was significantly more common in bronchoscopy controls than cases (p = 0.05). Bronchiectasis was found significantly more frequently in both control groups than cases; pulmonary fibrosis was a rare diagnosis in all groups (Table 3.6).

3.2.3 Non-Respiratory Disease

A positive non-respiratory medical history, including diagnoses such as ischaemic heart disease, hypertension and diabetes mellitus was reported in 25.6% of cases (n = 42), 27.2% of bronchoscopy controls (n = 59) and 47.8% of chest clinic controls (n = 43) (Table 3.7). A positive medical history was therefore significantly more frequent in chest clinic controls (p = 0.0003), but not in bronchoscopy controls (p = 0.73) when compared to cases. The distribution of the five most common benign non-respiratory diagnoses is detailed in table 3.7. There was no significant difference between cases and bronchoscopy controls for any specific diagnosis. Chest clinic controls were diagnosed with ischaemic heart disease (p = 0.04) and hypertension (0.004) significantly more frequently than cases.

Table 3.6 The distribution of chronic respiratory disease in the study population.

Chronic Respiratory Diagnoses	Cases	Controls			
		Bronchoscopy		Chest Clinic	
	n (%)	n (%)	p value[a]	n (%)	p value[a]
Any	63 (38.4)	112 (51.6)	**0.01**	79 (87.8)	**0.0003**
COPD / Emphysema	45 (27.4)	65 (30.0)	0.59	72 (80.0)	**0.0009**
Asthma	10 (6.1)	26 (12.0)	**0.05**	9 (10.0)	0.26
Bronchiectasis	0	8 (3.7)	**0.01**	7 (7.8)	**0.001**
Pulmonary Fibrosis	3 (1.8)	3 (1.4)	-	2 (2.2)	-

[a]comparison with cases

Table 3.7 The distribution of chronic non-respiratory disease in the study population.

Chronic Non-Respiratory Diagnoses	Cases	Controls			
		Bronchoscopy		Chest Clinic	
	n (%)	n (%)	p value[a]	n (%)	p value[a]
Any	42 (25.6)	59 (27.2)	0.73	43 (47.8)	**0.0003**
Type 2 Diabetes	10 (6.1)	8 (3.7)	0.27	10 (11.1)	0.16
Ischaemic Heart Disease	15 (9.1)	23 (10.6)	0.64	20 (22.2)	**0.04**
Cardiac Failure	3 (1.8)	3 (1.4)	0.73	4 (4.4)	0.22
Hypertension	5 (3.0)	15 (6.9)	0.09	11 (12.2)	**0.004**
Other Vascular	10 (6.1)	9 (4.1)	0.39	8 (8.9)	0.41

[a]comparison with cases

3.3 Demographic Details of Cases and Controls

3.3.1 Gender and Ethnicity

All of the study population was of Caucasian origin. Approximately two-thirds of cases were male (67.7%, n = 111) and one-third female (32.3%, n = 53). There was no significant difference in gender distribution between cases and either control group (Table 3.8).

Table 3.8 The gender of cases and controls.

Gender	Cases n (%)	Controls			
		Bronchoscopy		Chest Clinic	
		n (%)	p value[a]	n (%)	p value[a]
Male	111 (67.7)	130 (59.9)	0.12	56 (62.2)	0.38
Female	53 (32.3)	87 (40.1)		34 (37.8)	

[a]comparison with cases

3.3.2 Age

The mean age of cases was 69.2 ± 10.3 years, this was significantly greater than the mean age of bronchoscopy controls (p = 0.0001) but not chest clinic controls (p = 0.21) (Table 3.9). The age of cases ranged from 42 to 96 years; during the recruitment period no individual with an incident diagnosis of lung cancer was excluded because of age criteria. The age distribution for cases and controls is illustrated in figure 3.3.

Table 3.9 Mean age and age range of cases and controls.

Age	Cases	Controls			
		Bronchoscopy		Chest Clinic	
		value	p value[a]	value	p value[a]
Mean (years) ± sd.	69.2 ± 10.3	65.0 ± 10.5	**0.0001**	67.6 ± 9.9	0.21
Range (years)	42 - 96	41 - 87	-	46 - 87	-

[a]comparison with cases

a) Cases
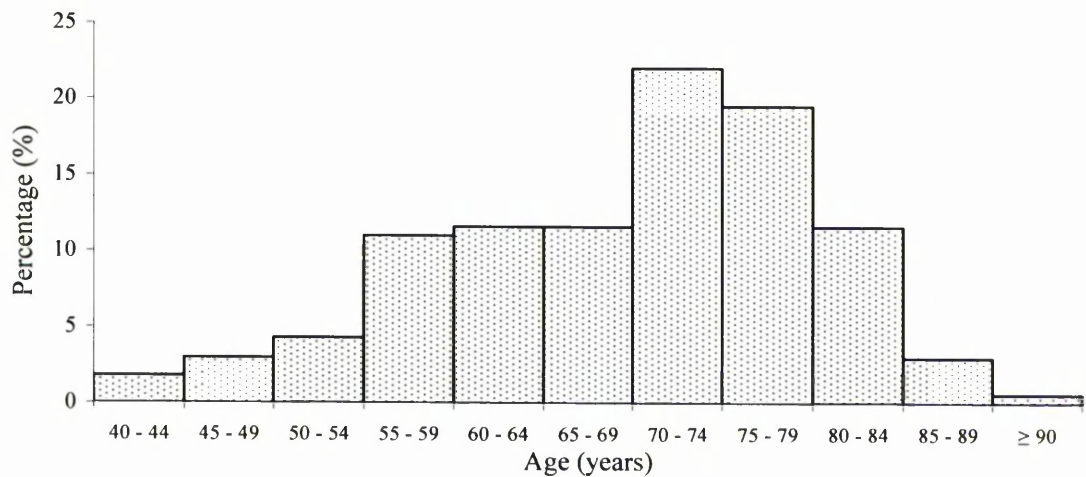
b) Bronchoscopy Controls
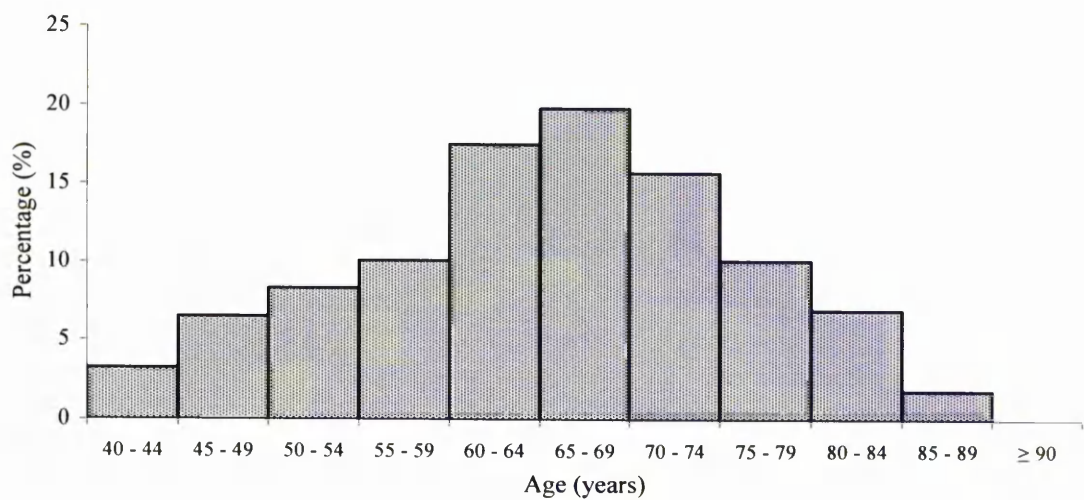
c) Chest Clinic Controls
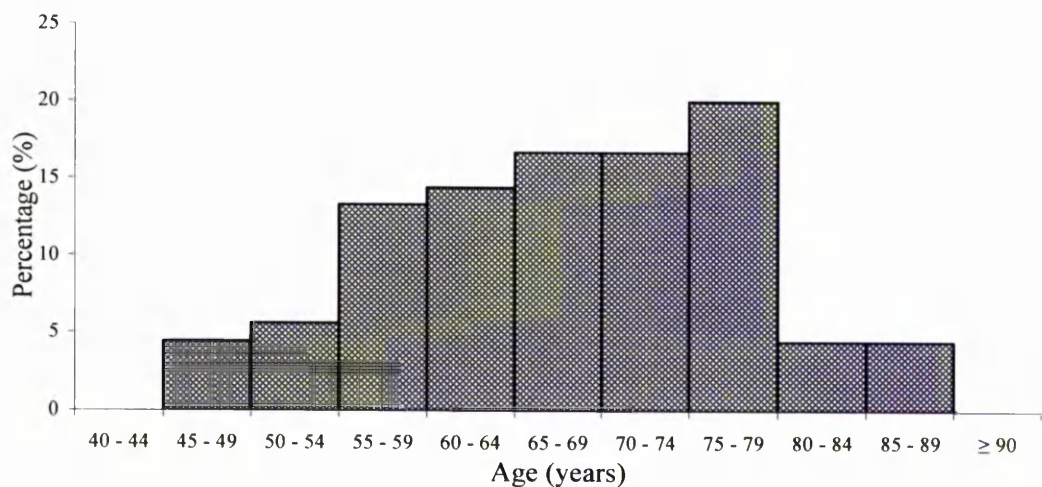
Figure 3.3 Age distribution of (a) cases, (b) bronchoscopy and (c) chest clinic controls.

3.4.3 Area, Duration of Residence and Deprivation

Cases had resided at their current address for a mean of 22.4 ± 16.5 years compared to 19.8 ± 14.7 years ($p = 0.12$) for bronchoscopy controls and 20.7 ± 16.1 years ($p = 0.44$) for chest clinic controls. The residential origin of cases and controls was analysed by comparing the most common postcode areas, which accounted for 90% of cases and 84% of controls (Figure 3.4). The graphical representation of postcode area suggested that the origin of subjects was similar; there was also no significant difference between cases and bronchoscopy controls ($p = 0.67$) or chest clinic controls ($p = 0.10$).
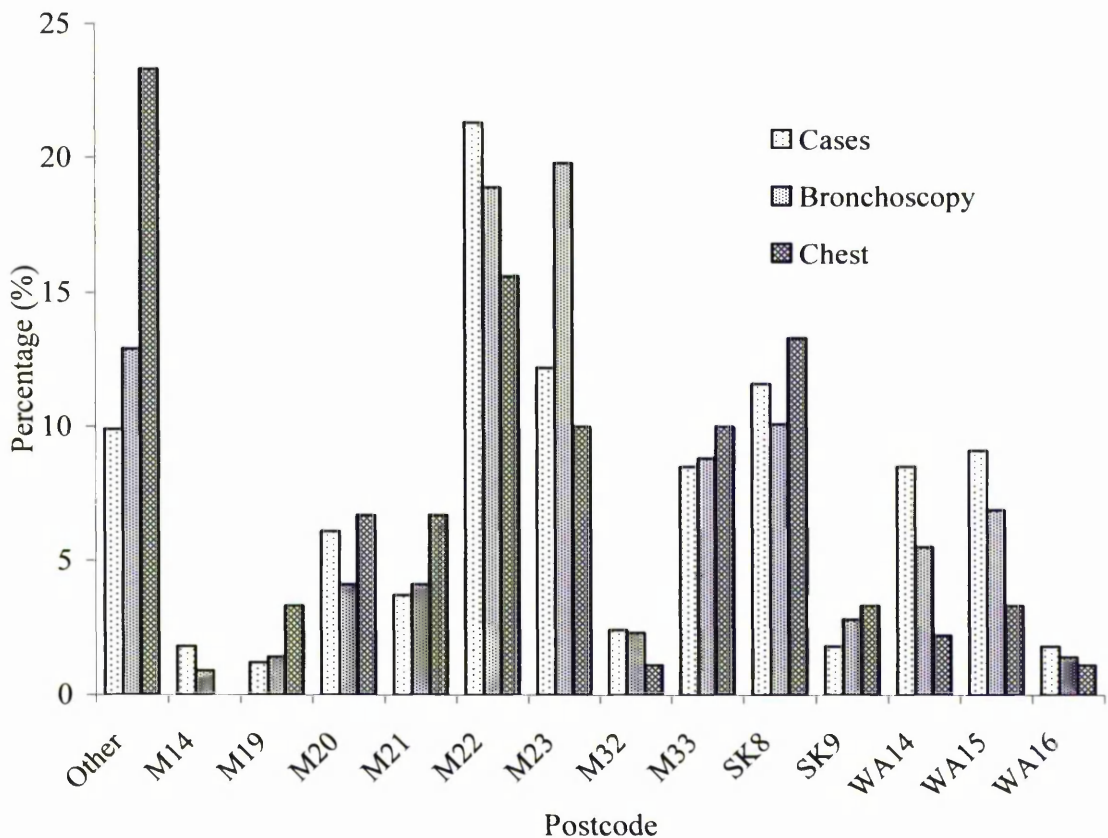
Figure 3.4 Postcode origins of cases and controls.

The IMD2004 ranking of each postcode was used as an indicator of deprivation (see Section 2.2.9). The rankings of all subjects spanned the full range of deprivation from 44 to 32230. The mean ranking of cases was higher than bronchoscopy controls (12343 ± 9927 vs. 10597 ± 9376), a difference that approached significance ($p = 0.08$). There was no significant difference between cases and mean ranking of chest clinic controls (11122 ± 8928, $p = 0.36$).

3.4.4 Tobacco Smoke Exposure

3.4.4.1 Smoking Status

The majority of subjects reported a positive smoking history; ever smokers represented 96.3% of cases, 89.9% of bronchoscopy controls and 97.8% of chest clinic controls (Table 3.10). Ever smoking was significantly more common in cases than bronchoscopy controls (p = 0.02) but not chest clinic controls (p = 0.53). Analysis of current smoking status in ever smokers revealed a greater proportion of former smokers to current smokers in chest clinic controls (p = 0.05) but a similar proportion in bronchoscopy controls (p = 0.83) when compared to cases (Table 3.10).

Table 3.10 Study subject smoking status.

Smoking Status	Cases n (%)	Controls			
		Bronchoscopy		Chest Clinic	
		n (%)	p value	n (%)	p value
Ever	158 (96.3)	195 (89.9)	**0.02**	88 (97.8)	0.53
Never	6 (3.7)	22 (10.1)		2 (2.2)	
Current[a]	65 (41.1)	78 (40.0)	0.83	25 (28.4)	**0.05**
Former[a]	93 (58.9)	117 (60.0)		63 (71.6)	

[a]values expressed as percentages of ever smokers only

3.4.4.2 Tobacco Product Smoked

Manufactured cigarettes were the predominant tobacco product, smoked by 91% (n = 401 / 441) of ever smoking study participants. The smoking of handrolled cigarettes accounted for 7.5% of smokers (n = 33 / 441). Only a small fraction of ever smokers smoked a pipe or cigars on a regular basis (n = 7 / 441, 1.5%). There was no difference in types of tobacco products smoked between the study groups (p = 0.74 and 0.23 for bronchoscopy and chest clinic controls respectively).

3.4.4.3 Extent of Tobacco Smoke Exposure

Cases started smoking tobacco at a mean age of 16.0 ± 4.4 years (range 7 to 50 years); although this was a younger age than either control group, the difference did not reach significance (Table 3.11). The mean duration of smoking tobacco was 46.3 ± 11.8 years in cases, this was on average 7.3 years longer than bronchoscopy controls (p = 0.0002) and 4.7 years longer than chest clinic controls (p = 0.004). Mean packyears exposure was greatest in chest clinic controls (59.6 ± 40.4 packyears) followed by cases (53.4 ± 27.8 packyears) and then bronchoscopy controls (47.2 ± 34.7 packyears). The difference between cases and chest clinic controls was not significant (p = 0.20), the difference between cases and bronchoscopy controls approached significance (p = 0.07) (Table 3.11).

Table 3.11 Tobacco exposure in ever smoking cases and controls.

Measures of tobacco smoke exposure in ever smokers		Cases	Controls			
			Bronchoscopy		Chest Clinic	
			value	p value[a]	value	p value[a]
Age Started Smoking	Mean (years) ± sd	16.0 ± 4.4	16.8 ± 4.2	0.11	16.7 ± 4.1	0.24
	Range (years)	7 - 50	5 - 35	-	8 - 35	-
Smoking Duration	Mean (years) ± sd	46.3 ± 11.8	39.0 ± 14.1	**0.0002**	41.6 ± 12.8	**0.004**
	Range (years)	11 - 77	3 - 67	-	7 - 71	-
Packyears	Mean (years) ± sd	53.4 ± 27.8	47.2 ± 34.7	0.07	59.6 ± 40.4	0.20
	Range (years)	0.7-138	0.4-177	-	1-161	-

[a]comparison with cases

The measure of tobacco exposure most strongly associated with lung cancer was therefore the duration of smoking.

3.4.4.4 Time from Smoking Starting to Lung Cancer Diagnosis

The mean length of time from first cigarette to diagnosis of lung cancer in ever smoking cases was 52.7 ± 10.1 years (range 25 to 77 years); this time was significantly shorter in current smokers (50.7 ± 10.5 years) compared to former smokers (54.1 ± 9.6 years, p = 0.04). The mean age of lifelong non-smokers at lung cancer diagnosis was 79.1 ± 10.5 years; this was significantly higher than ever smoking cases at diagnosis (68.8 ± 10.2 years, p = 0.02) and approaching significance when compared with non-smoking bronchoscopy controls (68.9 ± 11.3, p = 0.06). No comparison was made to chest clinic controls because only two subjects were non-smokers.

3.4.5   Recent Alcohol Intake

Subjects were asked to estimate their weekly alcohol intake over the preceding three months. A large proportion of each group had not taken any alcohol during this time period; the proportion of subjects who had not drunk alcohol was similar between cases and controls (Table 3.12).

Table 3.12 Recent weekly alcohol intake in cases and controls.

Alcohol	Cases	Controls			
		Bronchoscopy		Chest Clinic	
		value	p value[a]	value	p value[a]
None vs. ≥ 1 units/week n (% none)	65 / 99 (39.6)	74 / 143 (34.1)	0.27	39 / 51 (43.3)	0.57
Weekly Intake Geometric Mean[b] $\pm$ sd.	13.2 ± 3.2	11.2 ± 3.5	0.28	8.1 ± 3.2	**0.02**
Range (units/week)	0 - 300	0 - 210	-	0 - 140	-

[a]comparison with cases

[b]mean alcohol intake of subjects with a positive alcohol history only

The geometric mean of alcohol intake was calculated for subjects who had drunk alcohol in the preceding three months. Chest clinic controls drank significantly less alcohol than cases (p = 0.02), but there was no significant difference in alcohol intake between cases and bronchoscopy controls (p = 0.28).

3.4.6 Occupational History

3.4.6.1 Work Status

The majority of study subjects were retired; this equated to 77.8% of chest clinic controls, followed by 72.0% of cases and 62.7% of bronchoscopy controls. The proportion of cases in current work, either full or part time, was significantly greater than chest clinic controls (21.3% vs. 8.9%) (p = 0.01), but not significantly different to bronchoscopy controls (25.4%) (p = 0.40) (Table 3.13).

Table 3.13 The current work status of cases and controls.

Current Work Status		Cases n (%)	Controls	
			Bronchoscopy n (%)	Chest Clinic n (%)
Working	Full Time	30 (18.3)	44 (20.3)	5 (5.6)
	Part Time	5 (3.0)	11 (5.1)	3 (3.3)
Not Working	Retired	118 (72.0)	136 (62.7)	70 (77.8)
	Ill Health	6 (3.7)	19 (8.8)	12 (13.3)
	Unemployed	3 (1.8)	2 (0.9)	0
	Other	2 (1.2)	5 (2.3)	0

The proportion of controls of working age who were not working because of ill health ranged from 8.8% in bronchoscopy controls to 13.3% in chest clinic controls, this compared to 3.7% of cases. This observation was consistent with previous findings of increased chronic respiratory disease in the control groups compared to the cases.

3.4.6.2 Categorisation of Current or Last Occupation

Study subjects were categorised according to current occupation if working or last reported occupation if retired or not working for other reasons. The categorisation of subjects was based on the SOC2000 major occupation categories. The distribution of occupational types did not significantly differ according to case status (Table 3.14) (Figure 3.5).

Table 3.14 SOC2000 occupation category for cases and controls.

SOC2000 Major Occupational Categories		Cases	Controls			
			Bronchoscopy		Chest Clinic	
		n (%)	n (%)	p value[a]	n (%)	p value[a]
1	Managers and senior officials	13 (8.0)	17 (7.9)	0.97	8 (9.8)	0.64
2	Professional occupations	11 (6.7)	14 (6.5)	0.92	5 (6.1)	0.85
3	Associate professional and technical occupations	13 (8.0)	15 (6.9)	0.70	2 (2.4)	0.09
4	Administrative and secretarial occupations	18 (11.0)	22 (10.2)	0.79	7 (8.5)	0.54
5	Skilled trades occupations	31 (19.0)	42 (19.4)	0.92	12 (14.6)	0.40
6	Personal service occupations	7 (4.3)	16 (7.4)	0.21	4 (4.9)	0.84
7	Sales and customer service occupations	11 (6.7)	12 (5.6)	0.63	9 (11.0)	0.25
8	Process, plant and machine operatives	30 (18.4)	34 (15.7)	0.49	15 (18.3)	0.98
9	Elementary occupations	29 (17.8)	44 (20.4)	0.53	20 (24.4)	0.22

[a]comparison with cases
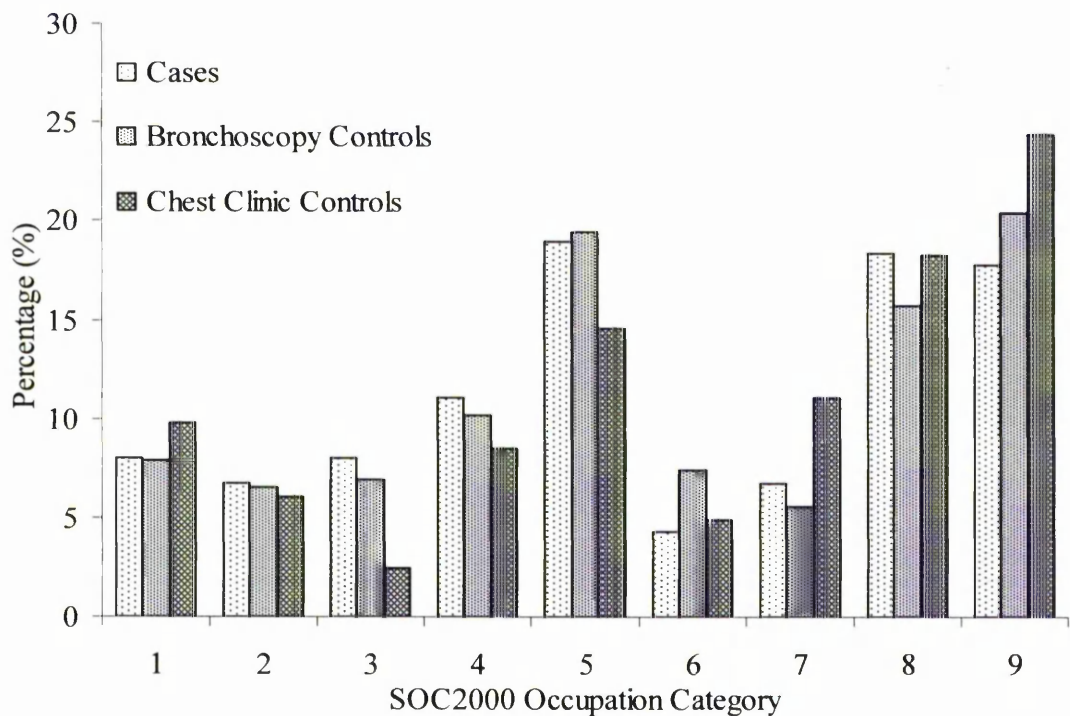
Figure 3.5 Distribution of occupational types according to case status

The most common occupation categories, accounting for over 50% of subjects in each study group were skilled trades, process/plant/machine operatives and elementary occupations. It was anticipated that the occupations of males and females would be significantly different therefore occupational data was subdivided according to gender (Table 3.15). Females were more likely to work in administrative/secretarial (p = 0.0004), personal service (p = 0.0002) and sales/customer service occupations (p = 0.006). Whereas males were more likely to work in skilled trades (p = 0.0003) or as process/plant/machine operatives (p = 0.0001).

Table 3.15 Comparison of study subject occupation according to gender.

	SOC2000 Major Occupation Categories	Male n (%)	Female n (%)	p value
1	Managers and senior officials	29 (9.1)	9 (5.3)	0.08
2	Professional occupations	21 (8.2)	9 (5.3)	0.42
3	Associate professional and technical occupations	15 (8.2)	15 (8.8)	0.12
4	Administrative and secretarial occupations	9 (2.7)	38 (22.4)	**0.0004**
5	Skilled trades occupations	76 (26.4)	9 (5.3)	**0.0003**
6	Personal service occupations	8 (2.7)	19 (11.2)	**0.0002**
7	Sales and customer service occupations	13 (2.7)	19 (11.2)	**0.006**
8	Process, plant and machine operatives	67 (23.6)	12 (7.1)	**0.0001**
9	Elementary occupations	53 (16.4)	40 (23.5)	0.17

The occupations of study participants were compared to the occupation distribution of a large sample from the general population of the UK, stratified by gender. This was designed to investigate whether the occupations and therefore indirectly the occupational exposures of study subjects, both cases and controls, were skewed because of the high prevalence of both benign and malignant respiratory disease (Table 3.16). There were significant or approaching significant differences in the proportions of male study subjects in eight out of nine occupational categories compared to the UK sample (Table 3.16) (Figure 3.6a). Male subjects were significantly more likely to be employed as managers

(OR 2.56, 95% CI 1.71-3.82), professionals (OR 1.68, 95% CI 1.05-2.66) or process, plant and machine operatives (OR 3.61, 95% CI 2.72-4.78) as the general male population. Female study subjects were significantly more likely to work in skilled trade occupations (OR 4.63, 95% CI 2.21-9.32) and as process, plant and machine operatives (OR 6.17, 95% CI 3.27-11.38). The proportion were lower in personal service (OR 0.60, 95% CI 0.36-0.99) and sales and customer service occupations (OR 0.35, 95% CI 0.21-0.58) (Figure 3.6b).

Table 3.16 Comparison of the occupations of study subjects to UK national statistics.

SOC2000 Major Occupation Categories	Male			Female		
	Study n (%)	UK n (%)	OR[a] (95% CI)	Study n (%)	UK n (%)	OR[a] (95% CI)
Managers and senior officials	29 (9.1)	82809 (4.1)	**2.56 (1.71-3.82)**	9 (5.3)	65288 (3.4)	1.58 (0.76-3.19)
Professional occupations	21 (8.2)	88530 (4.4)	**1.68 (1.05-2.66)**	9 (5.3)	80075 (4.2)	1.28 (0.61-2.58)
Associate professional and technical occupations	15 (5.2)	198465 (9.9)	**0.48 (0.27-0.82)**	15 (8.8)	171061 (8.9)	0.99 (0.56-1.71)
Administrative and secretarial occupations	9 (2.7)	154684 (7.7)	**0.38 (0.19-0.77)**	38 (22.4)	367517 (19.2)	1.21 (0.83-1.76)
Skilled trades occupations	76 (26.4)	438200 (21.9)	1.26 (0.96-1.65)	9 (5.3)	22846 (1.2)	**4.63 (2.21-9.32)**
Personal service occupations	8 (2.7)	61182 (3.1)	0.90 (0.41-1.86)	19 (11.2)	329689 (17.2)	**0.60 (0.36-0.99)**
Sales and customer service occupations	13 (2.7)	339096 (16.9)	**0.23 (0.13-0.41)**	19 (11.2)	505477 (26.4)	**0.35 (0.21-0.58)**
Process, plant and machine operatives	67 (23.6)	153135 (7.7)	**3.61 (2.72-4.78)**	12 (7.1)	23264 (1.2)	**6.17 (3.27-11.4)**
Elementary occupations	53 (16.4)	484976 (24.7)	**0.70 (0.51-0.95)**	40 (23.5)	348104 (18.2)	1.38 (0.96-2.00)

[a]odds of being a study participant

The differences seen in occupation between the study and UK populations could be due to the influence of selection bias. Subjects were selected from a population with a high prevalence of respiratory disease and would therefore be more likely to have undertaken occupations that may have had a role in causing or exacerbating respiratory conditions. However, some differences may also be due to fundamental differences in population demographics. The majority of study subjects were retired, the last job was more likely to be in a managerial post compared to a first job. Changes may also have occurred in the types of occupations undertaken over time, for instance less process, plant and machine operatives and more sales and customer service occupations.

(a) Males
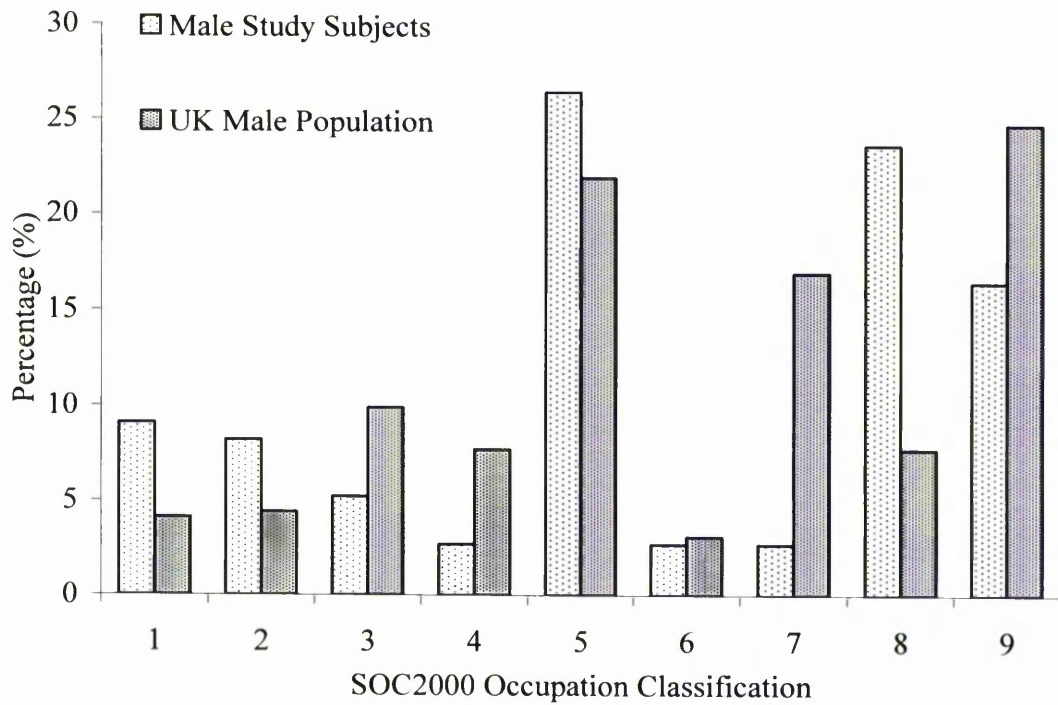
(b) Females
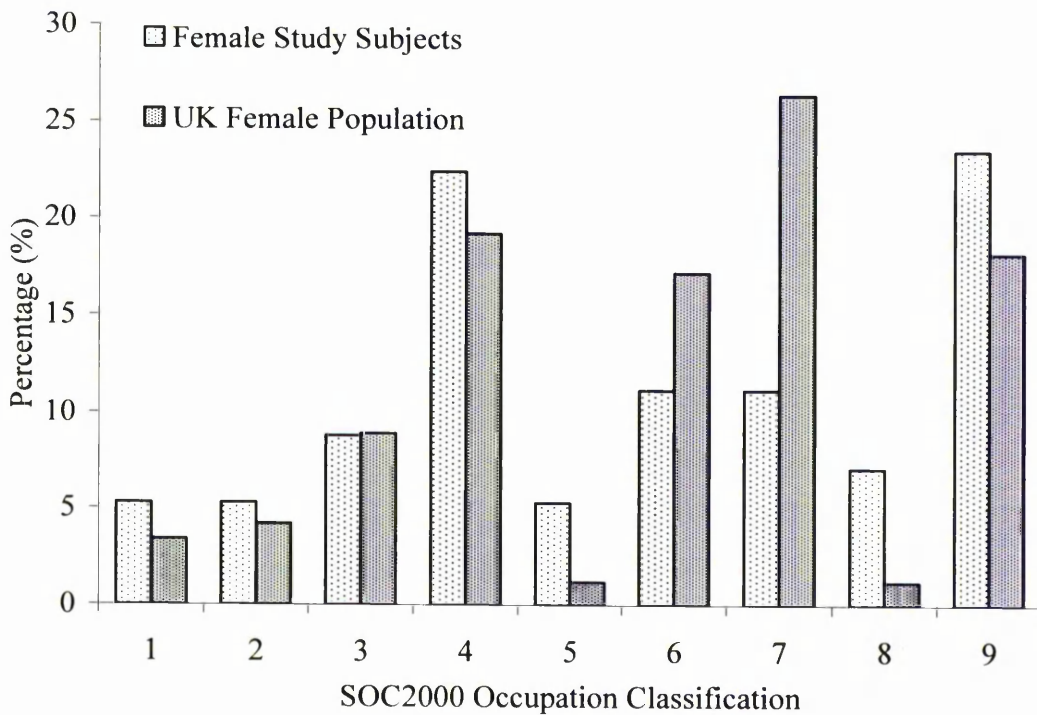
Figure 3.6 Distribution of SOC2000 occupation classification in (a) male and (b) female study subjects compared to UK national statistics (2005).

Subjects were asked to identify whether they had ever experienced significant exposure to certain substances during their occupational lifetime (Table 3.17) (Figure 3.7). The majority of cases reported at least one exposure (71.3%, n = 117 / 164), a similar proportion of bronchoscopy controls (72.4%, p = 0.83) and an even greater proportion of chest clinic controls (84.4%, p = 0.02) also reported a positive occupational history. The most commonly reported exposure was to dust, reported by 51.8% (n = 85) of cases, 58.5% (n = 127, p = 0.19) of bronchoscopy controls and 70.0% (n = 63) of chest clinic controls, which was significantly greater than cases (p = 0.005).

Table 3.17 Reported ever occupational exposure in cases and controls.

	Reported Ever Occupational Exposure	**Cases** n (%)	**Bronchoscopy Controls**		**Chest Clinic Controls**	
			n (%)	p value	n (%)	p value
1	Any	117 (71.3)	157 (72.4)	0.83	76 (84.4)	**0.02**
2	Dust	85 (51.8)	127 (58.5)	0.19	63 (70.0)	**0.005**
3	Fumes	60 (36.6)	87 (40.1)	0.49	43 (47.8)	0.08
4	Oils	29 (17.7)	58 (26.7)	**0.04**	14 (15.6)	0.67
5	Plastics/synthetic rubbers	23 (14.0)	44 (20.3)	0.11	20 (22.2)	0.10
6	Asbestos	38 (23.2)	48 (22.1)	0.81	28 (31.1)	0.17
7	Pesticides	1 (0.6)	8 (3.7)	**0.05**	5 (5.6)	**0.01**
8	Wood preservatives	12 (7.3)	20 (9.2)	0.51	9 (10.0)	0.46
9	Sealants	23 (14.0)	27 (12.4)	0.65	12 (13.3)	0.88
10	Acids	7 (4.3)	21 (9.7)	**0.05**	12 (13.3)	**0.009**
11	Adhesives/glues	21 (12.8)	39 (18.0)	0.17	18 (20.0)	0.13
12	Cleaning agents	36 (22.0)	72 (33.2)	**0.02**	36 (40.0)	**0.002**
13	Disinfecting agents	20 (12.2)	41 (18.9)	0.08	26 (28.9)	**0.001**
14	Ionising radiation or radioactive materials	4 (2.4)	2 (0.9)	0.24	3 (3.3)	0.68

Reported exposure was significantly more frequent to pesticides, acids and cleaning agents in both control groups than cases. Exposure to disinfecting agents was also more frequently reported in chest clinic controls and to oils in bronchoscopy controls compared to cases. Exposure to asbestos, a known occupational lung carcinogen, was reported in 23.2% (n = 38) of cases this was not significantly different to reported exposure in bronchoscopy (22.1%, n = 48, p = 0.81) or chest clinic controls (31.1%, n = 28, p = 0.17). Out of thirteen occupational exposures detailed on the risk factor questionnaire, nine were reported more frequently in controls, no exposure was reported more frequently by cases. The occupational differences according to gender were reflected in the very significant differences in reported occupational exposures (Table 3.18).

Table 3.18 Comparison of reported ever occupational exposure according to gender.

	Reported Ever Occupational Exposure	Male n (%)	Female n (%)	p value
1	Any	241 (81.1)	109 (62.6)	0.0009
2	Dust	195 (65.7)	80 (46.0)	0.0003
3	Fumes	158 (53.2)	32 (18.4)	0.0001
4	Oils	82 (27.6)	19 (10.9)	0.0002
5	Plastics/synthetic rubbers	69 (23.2)	18 (10.3)	0.001
6	Asbestos	107 (36.0)	7 (4.0)	0.0005
7	Pesticides	11 (3.7)	3 (1.7)	0.22
8	Wood preservatives	40 (13.5)	1 (0.6)	0.0002
9	Sealants	60 (20.2)	2 (1.1)	0.0004
10	Acids	38 (12.2)	2 (1.1)	0.0001
11	Adhesives/glues	66 (22.2)	12 (6.9)	0.0002
12	Cleaning agents	86 (29.0)	58 (33.3)	0.32
13	Disinfecting agents	47 (15.8)	40 (23.0)	0.05
14	Ionising radiation or radioactive materials	9 (3.0)	0 (0)	0.02

Males reported significantly more occupational exposure overall and significantly more exposure to ten out of the thirteen substances detailed in the risk factor questionnaire. Reported ever occupational exposure was therefore analysed according to case status after stratification by gender (Table 3.19a+b) (Figure 3.8a+b).

Table 3.19a Reported ever occupational exposure in male cases and controls.

	Reported Ever Occupational Exposure	Cases n (%)	Bronchoscopy Controls		Chest Clinic Controls	
			n (%)	p value	n (%)	p value
1	Any	88 (79.3)	102 (78.5)	0.88	51 (91.1)	**0.05**
2	Dust	67 (60.4)	86 (66.2)	0.35	42 (75.0)	0.06
3	Fumes	53 (47.7)	71 (54.6)	0.29	34 (60.7)	0.11
4	Oils	26 (23.4)	45 (34.6)	0.06	11 (19.6)	0.58
5	Plastics/synthetic rubbers	21 (18.9)	33 (25.4)	0.23	15 (26.8)	0.24
6	Asbestos	35 (31.5)	46 (35.4)	0.53	26 (46.4)	0.06
7	Pesticides	0 (0)	7 (5.4)	**0.01**	4 (7.1)	**0.004**
8	Wood preservatives	12 (10.8)	19 (14.6)	0.38	9 (16.1)	0.33
9	Sealants	22 (19.8)	27 (20.8)	0.86	11 (19.6)	0.98
10	Acids	6 (5.4)	20 (15.4)	**0.01**	12 (21.4)	**0.002**
11	Adhesives/glues	19 (17.1)	31 (23.8)	0.20	16 (28.6)	0.09
12	Cleaning agents	28 (25.2)	39 (30.0)	0.41	19 (33.9)	0.24
13	Disinfecting agents	14 (12.6)	20 (15.4)	0.54	13 (23.2)	0.08
14	Ionising radiation or radioactive materials	4 (3.6)	2 (1.5)	0.31	3 (5.4)	0.59

Table 3.19b Reported ever occupational exposure in female cases and controls.

	Reported Ever Occupational Exposure	Cases n (%)	Bronchoscopy Controls		Chest Clinic Controls	
			n (%)	p value	n (%)	p value
1	Any	29 (54.7)	55 (63.2)	0.32	25 (73.5)	0.08
2	Dust	18 (34.0)	41 (47.1)	0.13	21 (61.8)	**0.01**
3	Fumes	7 (13.2)	16 (18.4)	0.42	9 (26.5)	0.12
4	Oils	3 (5.7)	13 (14.9)	0.09	3 (8.8)	0.57
5	Plastics/synthetic rubbers	2 (3.8)	11 (12.6)	0.08	5 (14.7)	0.07
6	Asbestos	3 (5.7)	2 (2.3)	0.30	2 (5.9)	0.97
7	Pesticides	1 (1.9)	1 (1.1)	0.72	1 (2.9)	0.75
8	Wood preservatives	0 (0)	1 (1.1)	0.43	0 (0)	-
9	Sealants	1 (1.9)	0 (0)	0.20	1 (2.9)	0.75
10	Acids	1 (1.9)	1 (1.1)	0.72	0 (0)	0.42
11	Adhesives/glues	2 (3.8)	8 (9.2)	0.23	2 (5.9)	0.65
12	Cleaning agents	8 (15.1)	33 (37.9)	**0.004**	17 (50.0)	**0.0004**
13	Disinfecting agents	6 (11.3)	21 (24.1)	0.06	13 (38.2)	**0.003**
14	Ionising radiation or radioactive materials	0 (0)	0 (0)	1	0 (0)	1
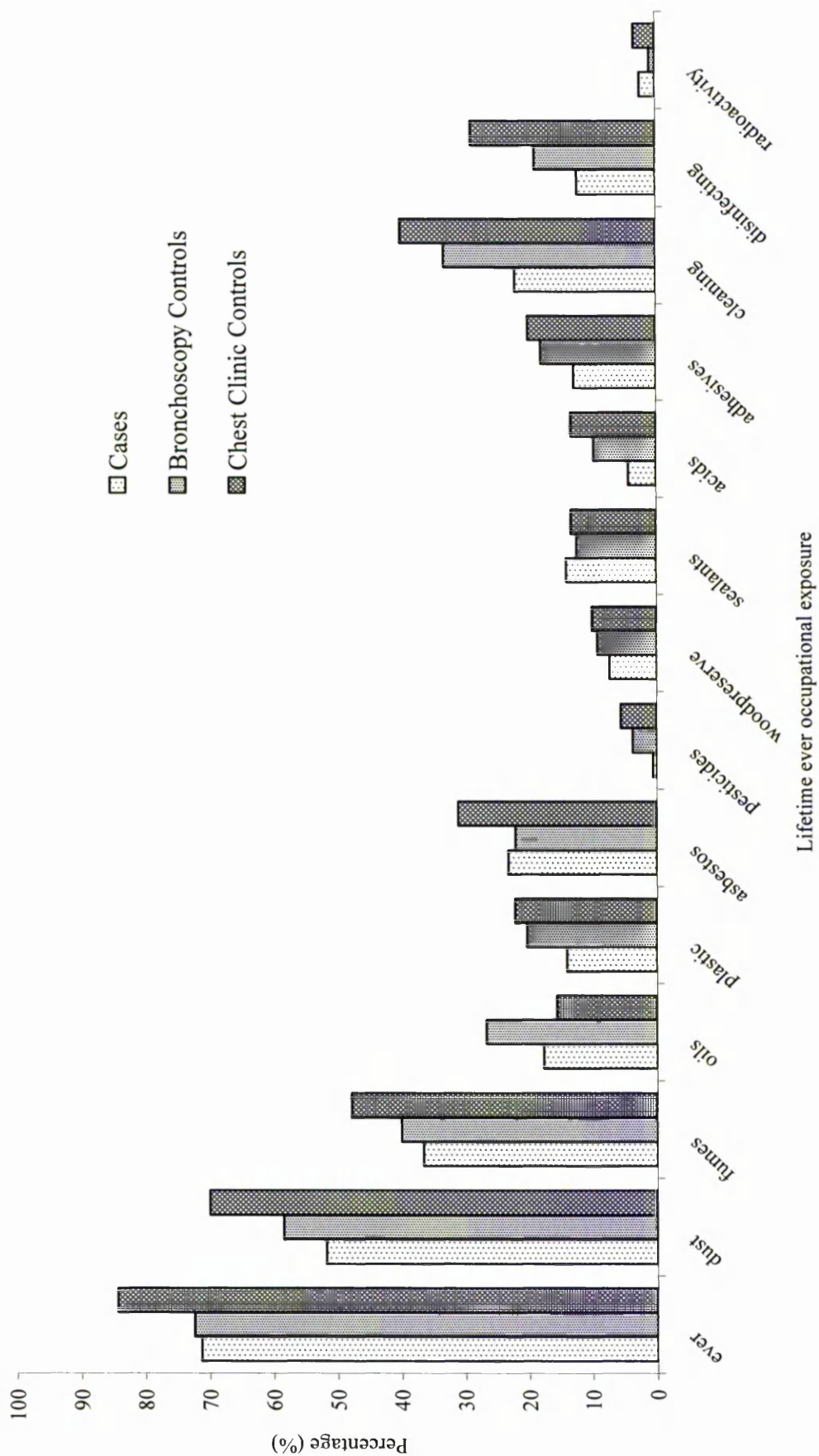
Figure 3.7 Lifetime ever occupational exposure in cases and controls.

112
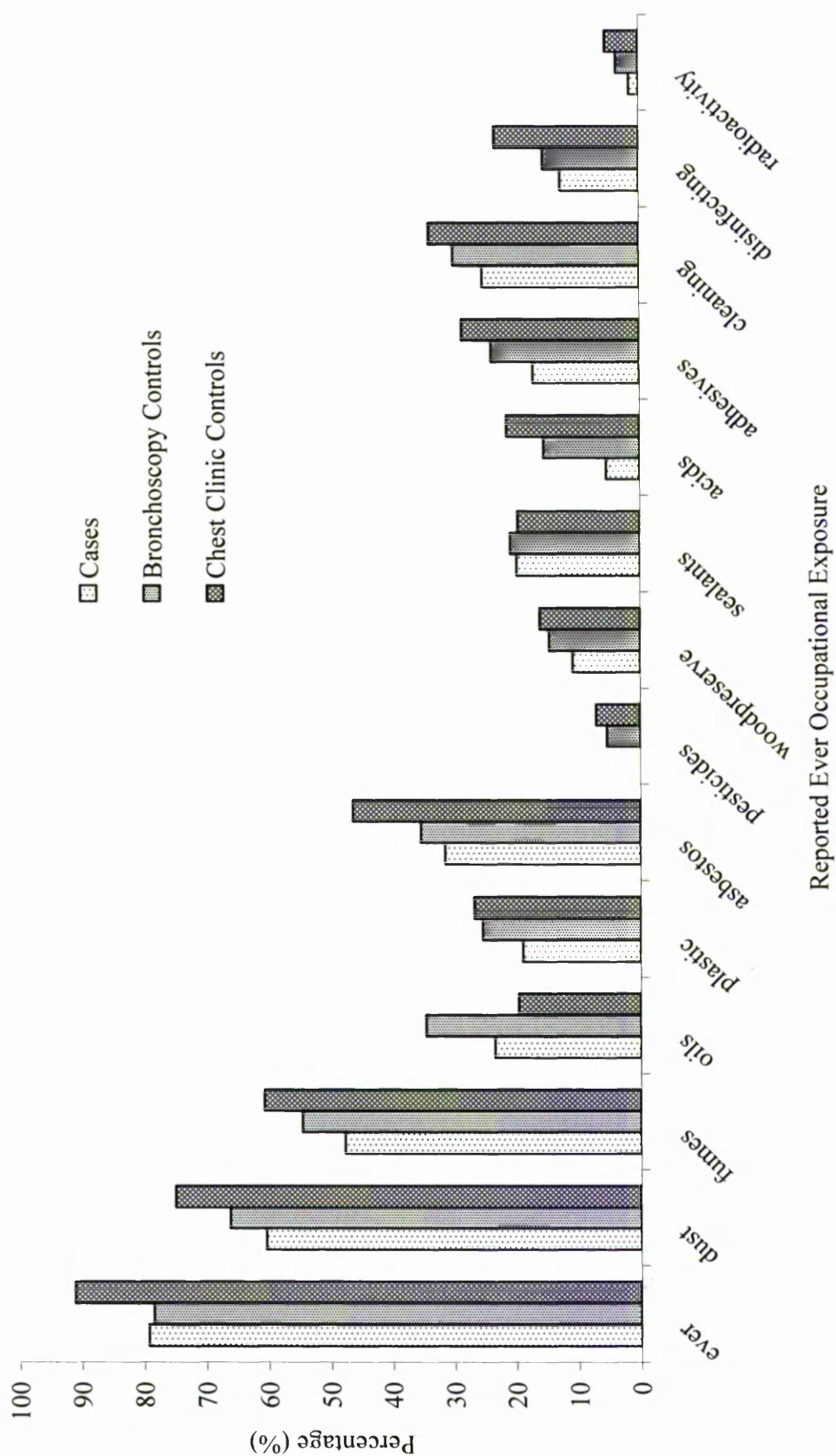
Figure 3.8a Ever reported occupational exposure in male cases and controls.

113
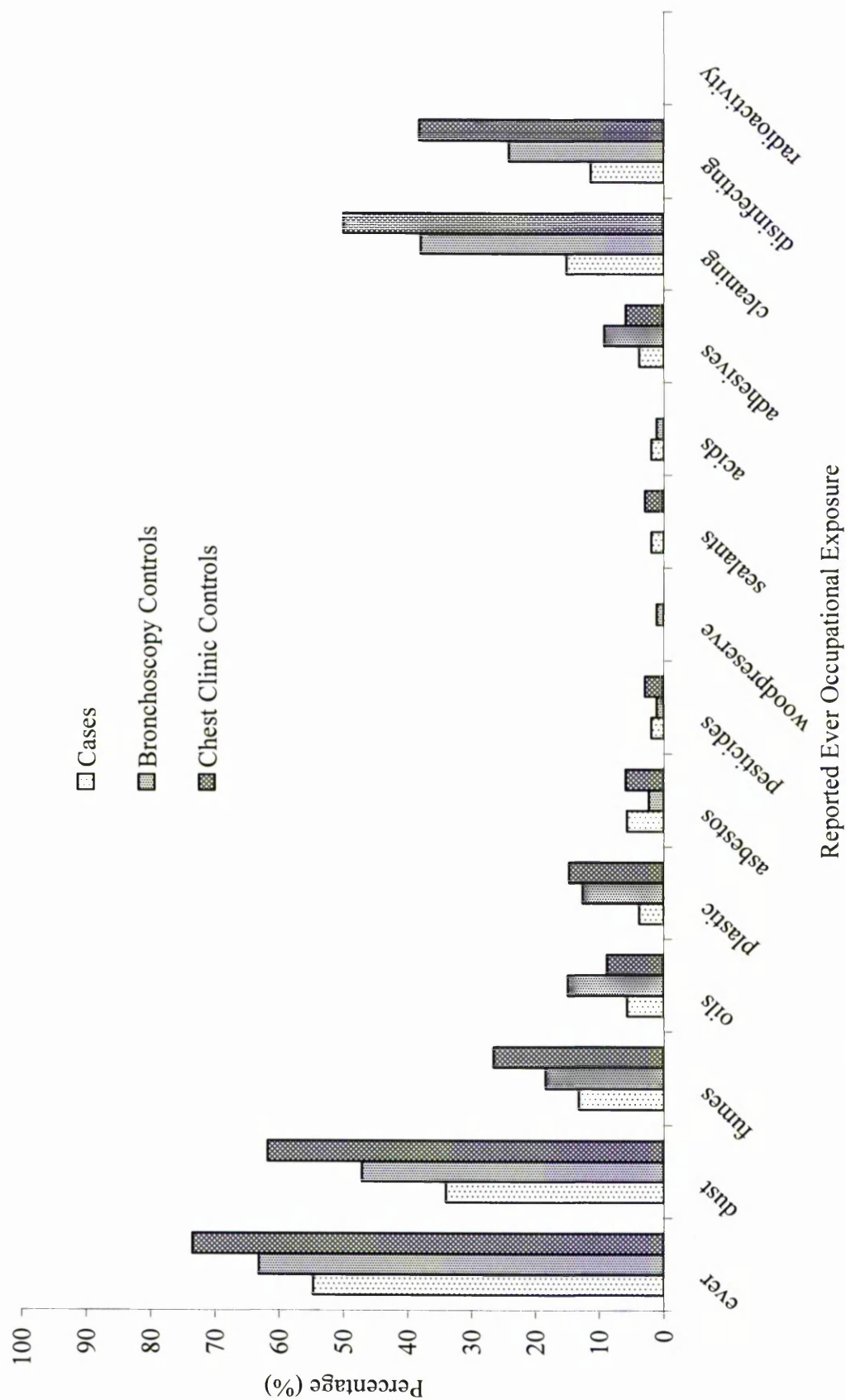
Figure 3.8b Ever reported occupational exposure in female cases and controls.

Reported Ever Occupational Exposure

114

3.4.7 Family History

3.4.7.1 Lung Cancer

One in five study participants reported a positive family history of lung cancer (the diagnosis of lung cancer in a first degree relative). There was no difference in family history of lung cancer between cases and either control group (Table 3.20).

Table 3.20 Family history of lung cancer in cases and controls.

Family History of Lung Cancer	Cases	Bronchoscopy Controls			Chest Clinic Controls		
	n (%)	n (%)	OR[a] (95% CI)	p value	n (%)	OR[a] (95% CI)	p value
No	129 (78.7)	167 (77.0)	1	0.67	72 (80.0)	1	0.74
Yes	34 (20.7)	49 (22.6)	0.90 (0.55–1.47)		17 (18.9)	1.12 (0.58-2.14)	
Unknown	1 (0.6)	1 (0.5)	-	-	1 (1.1)	-	-

[a]odds of lung cancer incidence (reference = no family history)

3.4.7.2 COPD

One third of cases and bronchoscopy controls reported a family history of COPD or 'wheezy bronchitis'. A significantly greater proportion of chest clinic controls reported a positive family history than cases (p = 0.03) (Table 3.21).

Table 3.21 Family history of COPD in cases and controls.

Family History of COPD	Cases	Bronchoscopy Controls			Chest Clinic Controls		
	n (%)	n (%)	OR[a] (95% CI)	p value	n (%)	OR[a] (95% CI)	p value
No	111 (67.7)	146 (67.3)	1	0.92	48 (53.3)	1	**0.03**
Yes	52 (31.7)	70 (32.3)	0.98 (0.63–1.51)		41 (45.6)	**0.55 (0.32–0.93)**	
Unknown	1 (0.6)	1 (0.5)	-	-	1 (1.1)	-	-

[a]odds of lung cancer incidence (reference = no family history)

3.5 Discussion

The aim of the study was to compare factors that might influence susceptibility to tobacco smoke carcinogens and therefore influence lung cancer risk. A case-control study design was chosen as this would allow the investigation of multiple factors with respect to lung cancer outcome. A crucial aspect of this study was the definition and identification of cases and controls.

The strategy for selecting cases was to recruit from the Bronchoscopy Unit. This area acted as a funnel through which all patients with a potential diagnosis of lung cancer were referred for investigation. The catchment area for this unit was the same as the catchment area of Wythenshawe Hospital, with the local populus contributing the vast majority of patients. A small subset of patients was not referred for bronchoscopy if they were not medically fit enough for the procedure. It was felt that this was an insignificant number and that overall patients with lung cancer investigated in the unit were representative of patients with lung cancer in the local community. Incident diagnoses of lung cancer were selected and prevalent disease excluded to ensure that cases were selected simply according to presence of disease and not according to the natural history or severity of the tumour. Patients with a previous cancer diagnosis were also excluded as the study was looking specifically at lung cancer susceptibility and not susceptibility to cancer as a whole. However, the exclusion of patients with a prior history of lung cancer may have introduced bias by removing the individuals most susceptible to tobacco smoke carcinogens.

Two control groups were chosen the first from the Bronchoscopy Unit and the second from the Chest Clinic. The bronchoscopy controls underwent the same selection process as subjects with lung cancer; in theory the only difference between the groups was the presence or absence of lung cancer. However, symptoms such as coughing up of blood, if not due to lung cancer could be due to underlying chronic respiratory conditions such as chronic obstructive pulmonary disease (COPD) or bronchiectasis or the acute presentation of a new condition such as pneumonia / bronchitis. The prevalence of chronic respiratory disease was high in all subjects recruited from the Bronchoscopy Unit; present in 51.6% of controls and 38.4% of subjects with lung cancer, a significant difference ($p = 0.01$).

However, the prevalence of diagnosed COPD the main tobacco related lung condition was not significantly different between the groups (p = 0.59), accounting for 58% and 71% of chronic lung disease diagnoses in bronchoscopy controls and cases respectively.

Subjects undergoing bronchoscopy were highly selected and it was therefore unclear how representative the findings of the study might be to a wider population. Since lung cancer is almost exclusively a disease of chronic smokers; the study needed to be representative of other chronic smokers and not to a wider healthy population not exposed to cigarette smoke. A second control group was chosen to reduce the possible impact of the selection process inherent in sampling subjects in the Bronchoscopy Unit. Patients attending the Chest Clinic were therefore randomly selected using the parameters set by the inclusion and exclusion criteria detailed in section 2.2.3. The vast majority of chest clinic controls had COPD (87.8%), which shares a common aetiological agent with lung cancer, namely chronic exposure to tobacco smoke. In a similar fashion to lung cancer only a fraction of chronic smokers develop COPD, suggesting a genetic predisposition (Barnes 2000). The chest clinic controls provided a hospital based population with chronic high tobacco exposure, who had not undergone the same selection procedure as bronchoscopy controls and so were more representative to a wider smoking population. Ideally, comparing cases to healthy elderly chronic smokers would have been better, but the recruitment of such a population would have been difficult in a hospital environment. Alternatives may have been recruitment from GP surgeries or recruitment from smoking cessation clinics.

Only Caucasian subjects were included in the study population because genotype distribution and lung cancer risk is known to differ between ethnic groups (Ingelman-Sundberg, Johansson et al. 1992; Garte, Gaspari et al. 2001; Weiserbs, Jacobson et al. 2003). The recruitment rate from subjects approached for study participation was high, 92% in the Bronchoscopy Unit and 95% from the Chest Clinic. The recruitment rate was similar for cases and controls (91% vs. 94%). Blood samples were taken from all study participants except for three bronchoscopy controls who were excluded from further analysis.

The collection of information was standardised as much as possible through the use of a standard risk factor questionnaire and the use of only one interviewer (PAJC). In the Bronchoscopy Unit the interview was conducted on the day of, but prior to the

bronchoscopy procedure to blind the investigator to the case status of the patient; so reducing observer bias. There was therefore no difference in data collection between cases and bronchoscopy controls. To a large degree patients were also blinded to their own case status, as the bronchoscopy was in the majority the first step in the diagnostic pathway after the initial Chest Clinic consultation. However, it was a possibility that the doctor undertaking the initial consultation may have forewarned some patients as to the likely outcome of their investigations.

A histological diagnosis was achieved in 80.5% of cases. The remaining cases were diagnosed after radiological and clinical review. In cases with no histology an abnormality was detected and sampled in 28% (n = 9 / 32) at bronchoscopy, suggesting false negative sampling. Because of the severity of illness of some patients and the invasive nature of the procedure, it was not always appropriate to repeat the bronchoscopy when a false negative result was suspected. However the reported appearance of tumour by the bronchoscopist increases the likelihood of lung cancer as a diagnosis in these cases even without histological confirmation. The histological split of cases and stage of disease at presentation is consistent with national statistics (Alberg and Samet 2003; Janssen-Heijnen and Coebergh 2003). The high level of smoking exposure in cases is also consistent with previous findings in the UK (Group 1987).

The comparison of tobacco smoke exposure between cases and controls showed that cases were more likely to have been ever smokers and to have smoked more over a longer period of time than bronchoscopy controls. Compared to chest clinic controls, cases were more likely to be current smokers and to have smoked over a longer period of time. The finding that smoking duration was greater in cases than either control group is consistent with reports suggesting that duration of smoking is a stronger risk factor than amount smoked for the development of lung cancer (Flanders, Lally et al. 2003). However, the use of self reported smoking history without objective validation using a biomarker such as cotinine may have resulted in misclassification of smoking status. This was illustrated in a previous study of this population which showed that 18% of self-reported current non-smokers had cotinine levels consistent with current smoking (Lewis, Cherry et al. 2003). It was decided not to perform cotinine measurements in this study because the aims were not directly related to smoking exposure.

Bronchoscopy controls had a lower mean tobacco exposure and were also significantly younger than cases, it might therefore be suggested that these controls could develop lung cancer in the future if left until the same mean age as cases. This is of course eminently possible. The one advantage this study has with respect to this problem are the intensive investigations subjects underwent prior to assignment of case status. Bronchoscopy controls have in a sense undergone CT and bronchoscopic screening, with both investigations needing to be negative before assignment of control status. All reasonable steps were therefore taken to exclude even the earliest forms of lung cancer in these subjects. Ideally, follow up of control patients would be carried out to determine if lung cancer does eventually develop in the future but this was not practical for this study.

Occupational history was taken from all participants because of the potential role of occupational exposures in the development of lung cancer. The current or last occupation categories of study subjects were used as a guide to the types of occupation undertaken. There were no significant differences in types of occupation between cases and controls before and after stratification for gender. However, there was some evidence that the types of occupations undertaken in the study group differed significantly from the UK population as a whole. For instance, there were more process, plant and machine operatives in the study population and more skilled trades undertaken by females than the UK population. Differences between the study population and the UK population may be due to a number of factors. Firstly the study population is in the main an older population with 68.8% (n = 324 / 471) of subjects retired and only 20.8% in current work (n = 98 / 471). The classification of occupations undertaken by subjects at the beginning of their career and at the end is often different, for example due to career progression and may be one reason why more male subjects were classified as managers. Secondly controls were selected from a population with high levels of chronic respiratory disease, as occupational exposure is known to cause or exacerbate respiratory conditions this may have skewed the occupations of both cases and controls.

The increased reporting of ever occupational exposures in the control groups, before and after stratification for gender, may well reflect the role of occupational exposure in the aetiology of benign chronic respiratory conditions or the exacerbation of pre-existing disease. Another possibility is that recall bias may have played a role. Subjects who had been diagnosed with a chronic respiratory problem in the past may have already discussed

possible occupational exposures as a cause of their condition. As chronic respiratory diagnoses were more common in controls this may have biased the responses to occupational exposure questions.

Overall, the control groups were comparable to cases in terms of age, gender, smoking and occupational histories. The use of two different control groups attempted to overcome some of the selection bias induced by hospital based recruitment. The collection of data by one individual using a standard questionnaire blinded to case status reduced observer and removed inter-observer bias. Differences in biological measures e.g. genotype or DNA repair activity were therefore more likely to reflect differences due to the presence or absence of lung cancer than the effect of confounding factors e.g. smoking.

Chapter Four

GSTM1, GSTT1, hOGG1, and MGMT Genotype Distribution in the Case-Control Study Population

4.1   GSTM1 Genotype Distribution

The GSTM1 homozygous gene deletion polymorphism was predicted to occur in 50% of the population overall and to increase the odds of lung cancer. Case-control studies examining the relationship of this null polymorphism with lung cancer number in excess of one hundred (Ye, Song et al. 2006). However, in all but one of these studies genotyping has been limited to the determination of the presence or absence of the GSTM1 gene. The aim therefore of this study was to use real time PCR relative quantitation techniques to determine the copy number of the GSTM1 gene in individuals who previously would have been classified as GSTM1 positive only. This would allow a more accurate assessment of the relationship between GSTM1 and lung cancer; it was hypothesised that subjects with two copies of GSTM1 might be at less of a risk of developing lung cancer than subjects with one copy and that reduction in risk would be greatest when compared to subjects null for GSTM1.

4.1.1   The Presence or Absence of GSTM1 and Lung Cancer

The presence or absence of the GSTM1 gene was successfully determined in 99.2% of subjects (n = 467 / 471); this equated to 97.6% of cases (160/164) and 100% of controls (n = 307 / 307). The four cases not genotyped included two subjects with squamous cell carcinoma, one with adenocarcinoma and one subject with no histology. In three cases PCR products were not detectable despite repeated assays and in one case the presence of an indeterminate line in the visualisation of the PCR product prevented accurate categorisation.

The GSTM1 homozygous null polymorphism occurred significantly more frequently in cases (n = 94 / 160, 58.8%) than either bronchoscopy (n = 104 / 217, 47.9%; $p = 0.04$) or chest clinic controls (n = 39 / 90, 43.3%; $p = 0.02$). Using subjects who had at least one copy of the GSTM1 gene as a reference, the adjusted odds ratio of lung cancer incidence was significantly increased in homozygous null individuals when cases were compared to bronchoscopy controls (OR_{adj} 1.57, 95% CI 1.02-2.43) and chest clinic controls (OR_{adj} 1.86, 95% CI 1.09-3.15) (Table 4.1).

Subset analysis of GSTM1 genotype distribution revealed a pattern of increased odds of lung cancer for the null genotype in all strata examined for bronchoscopy controls (Table 4.2). The trend was similar in chest clinic controls with significance reached for males (OR_{adj} 1.92, 95% CI 0.99-3.72, p = 0.05) and controls aged above the median age of 65.5 years (OR_{adj} 2.45, 95% CI 1.24-4.84, p = 0.01). This contrasted with the below median age category where no apparent effect was seen (OR_{adj} 1.13, 95% CI 0.47-2.70).

The stratification of cases according to histology revealed that the strongest association with the null genotype was with large cell carcinoma when compared with both bronchoscopy (OR_{adj} 4.80, 95% CI 1.44-16.0, p = 0.01) and chest clinic controls (OR_{adj} 3.99, 95% CI 1.26-12.6, p = 0.02) (Table 4.3). All NSCLC when analysed together, were also significantly associated with the null genotype in both bronchoscopy (OR_{adj} 1.70, 95% CI 1.02-2.85, p = 0.04) and chest clinic control group comparisons (OR_{adj} 2.05, 95% CI 1.12-3.73, p = 0.02).

Table 4.1 GSTM1 gene distribution in the study population.

	Cases	Controls					
		Bronchoscopy			Chest Clinic		
GSTM1	N (%)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
Present	66 (41.3)	113 (52.1)	1	1	51 (56.7)	1	1
Null	94 (58.8)	104 (47.9)	1.55 (1.03-2.34)	1.57 (1.02-2.43)	39 (43.3)	1.86 (1.11-3.14)	1.86 (1.09-3.15)

[a]odds of lung cancer incidence (reference = GSTM1 present); [b]adjusted for age, sex and smoking duration

124

Table 4.2 GSTM1 gene distribution in the study population after stratification for: age, gender and smoking duration.

Factor	Strata	Cases	Controls					
			Bronchoscopy			Chest clinic		
		N +/- (%)	N +/- (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)	N +/- (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)
Age	≤ 65.5	22 / 30 (42.3/57.7)	56 / 53 (51.4/48.6)	1.44 (0.74-2.81)	1.50[cd] (0.76-2.96)	17 / 19 (47.2/52.8)	1.22 (0.52-2.87)	1.13[cd] (0.47-2.70)
	> 65.5	44 / 64 (40.7/59.3)	57 / 51 (52.8/47.2)	1.63 (0.95-2.79)	1.62[cd] (0.92-2.84)	34 / 20 (63.0/37.0)	**2.47 (1.26-4.85)**	**2.45[cd] (1.24-4.84)**
Gender	Male	43 / 65 (39.8/60.2)	67 / 63 (51.5/48.5)	1.61 (0.96-2.70)	1.52[bd] (0.89-2.60)	32 / 24 (57.1/42.9)	**2.02 (1.05-3.88)**	**1.92[bd] (0.99-3.72)**
	Female	23 / 29 (44.2/55.8)	46 / 41 (52.9/47.1)	1.42 (0.71-2.82)	1.79[bd] (0.84-3.83)	19 / 15 (55.9/44.1)	1.60 (0.67-3.81)	1.68[bd] (0.69-4.13)
Smoking Duration	≤ 40 yrs	16 / 27 (37.2/62.8)	58 / 58 (50.0/50.0)	1.69 (0.82-3.46)	1.70[bc] (0.83-3.50)	20 / 15 (57.1/42.9)	2.25 (0.90-5.60)	2.28[bc] (0.91-5.71)
	> 40 yrs	50 / 67 (42.7/57.3)	55 / 46 (54.5/45.5)	1.60 (0.94-2.74)	1.57[bc] (0.91-2.72)	31 / 24 (56.4/43.6)	1.73 (0.91-3.30)	1.80[bc] (0.93-3.46)

[a]odds of lung cancer incidence (reference = GSTM1 present); OR adjusted for [b]age, [c]sex and [d]smoking duration

(GSTM1 present (+), GSTM1 null (-)).

Table 4.3 GSTM1 gene distribution in the study population after stratification of cases according to histology.

Strata	Cases N+/- (%)	Controls					
		Bronchoscopy			Chest clinic		
		N+/- (%)	Unadjusted[a] OR (95%CI)	Adjusted[ab] OR (95%CI)	N+/- (%)	Unadjusted[a] OR (95%CI)	Adjusted[ab] OR (95%CI)
Squamous Cell	16/25 (39.0/61.0)	113/104 (52.1/47.9)	1.70 (0.86-3.36)	1.58 (0.77-3.26)	51/39 (56.7/43.3)	2.04 (0.96-4.34)	1.73 (0.78-3.84)
Adenocarcinoma	16/18 (47.1/52.9)		1.22 (0.59-2.52)	1.23 (0.59-2.56)		1.47 (0.67-3.25)	1.51 (0.68-3.39)
Small Cell	15/16 (48.4/51.6)		1.16 (0.55-2.46)	1.24 (0.57-2.71)		1.40 (0.62-3.16)	1.35 (0.57-3.19)
Large Cell	4/14 (22.2/77.8)		**4.58 (1.40-15.0)**	**4.80 (1.44-16.0)**		**3.80 (1.21-11.9)**	**3.99 (1.26-12.6)**
All NSCLC	36/57 (38.7/61.3)	113/104 (52.1/47.9)	**1.72 (1.05-2.82)**	**1.70 (1.02-2.85)**	51/39 (56.7/43.3)	**2.07 (1.15-3.74)**	**2.05 (1.12-3.73)**

[a] odds of lung cancer incidence (reference = GSTM1 present); [b] OR adjusted for age, sex and smoking duration

126

4.1.2 GSTM1 Copy Number and Lung Cancer

At least one copy of the GSTM1 gene was detected in 49.3% (n = 230 / 467) of study subjects for whom initial GSTM1 genotyping was successfully achieved. The aim of the next aspect of this study was to determine in these individuals the number of copies of GSTM1 present using real time multiplex PCR techniques (detailed in section 2.4.11). Samples were run in triplicate and each sample was assayed on three separate occasions. A mean ΔC_T (±sd) was calculated for each sample. When a histogram of mean ΔC_T was plotted a bimodal distribution was seen (Figure 4.1).
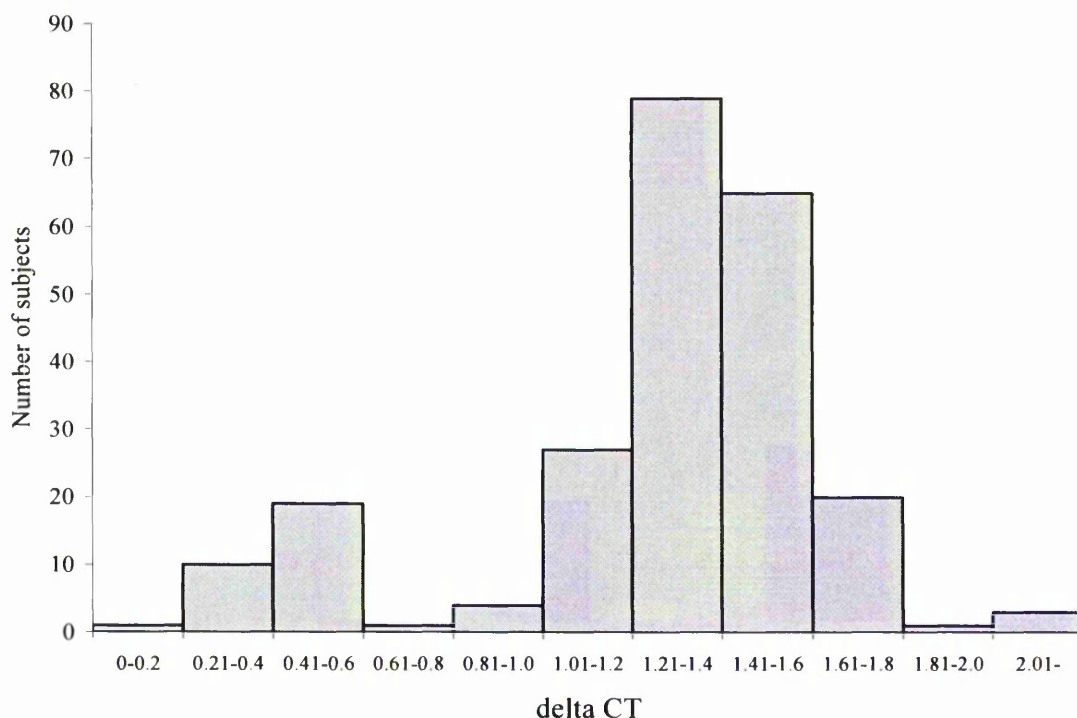
Figure 4.1 The frequency of ΔC_T values in subjects with at least one copy of GSTM1.

It was possible to divide the population according to mean ΔC_T into two groups (above and below mean ΔC_T 0.75) (Figure 4.2). Three subjects were not plotted on the graph as their mean ΔC_T was in excess of three (values 4.14, 3.46 and 5.64) and were therefore deemed to be outliers. The mean ΔC_T of the group above the dividing line was 1.38 ± 0.17 (n = 196) (excluding the 3 outliers), this was consistent with these subjects possessing one copy of the GSTM1 gene. The mean ΔC_T of the group below the dividing line was 0.45 ± 0.10 (n = 31). The mean ΔC_T of the smaller group was consistent with these subjects

having two GSTM1 gene copies. The difference of 0.93 between the groups was highly significant ($p = 0.0006$).

The mean $\Delta C_T \pm 3$ standard deviations were used to define the accepted range for each group and so define the accepted range for gene copy number. The ranges for one copy was 0.86 to 1.90 and for two copies 0.14 to 0.76; as there was no overlap between the groups it was possible to assign each subject a GSTM1 copy number. The three subjects with very high ΔC_T values were assigned to the one copy group for the purpose of statistical analysis. Overall, subjects with two copies of the GSTM1 gene represented 6.6% of the whole study population ($n = 31 / 467$) compared to 42.6% for subjects with one gene copy ($n = 199 / 467$). The GSTM1 allele was significantly less frequent in cases than chest clinic controls (0.23 vs. 0.33, $p = 0.02$); the difference with bronchoscopy controls approached significance (0.23 vs. 0.29, $p = 0.07$) (Table 4.4). Using the null genotype as a reference, the odds of lung cancer in subjects with one copy of the GSTM1 gene were significantly lower when cases were compared with bronchoscopy (OR_{adj} 0.63, 95% CI 0.40-0.99; $p = 0.05$) and chest clinic controls (OR_{adj} 0.56, 95% CI 0.32-0.97; $p = 0.04$) (Table 4.5). The odds of lung cancer were also reduced in subjects with two copies of GSTM1 but findings did not reach significance for either bronchoscopy (OR_{adj} 0.67, 95% CI 0.27-1.71; $p = 0.41$) or chest clinic controls (OR_{adj} 0.43, 95% CI 0.15-1.22; $p = 0.11$). This was possibly due to the low frequency of GSTM1 positive homozygotes in the study population. However, a trend test was significant when cases were compared to chest clinic controls ($p = 0.02$) and approached significance when compared to bronchoscopy controls ($p = 0.07$) (Table 4.5).

Table 4.4 Distribution of GSTM1 gene copy number.

Study Group	GSTM1 Copy Number N (%)			Total	Allele Frequency	p value
	2	1	None			
Cases	9 (5.6)	57 (35.6)	94 (58.8)	160	0.23	-
Bronchoscopy Controls	14 (6.5)	99 (45.6)	104 (47.9)	217	0.29	0.07
Chest Clinic Controls	8 (8.9)	43 (47.8)	39 (43.3)	90	0.33	**0.02**
Total	31 (6.6)	199 (42.6)	237 (50.7)	467	-	-

128
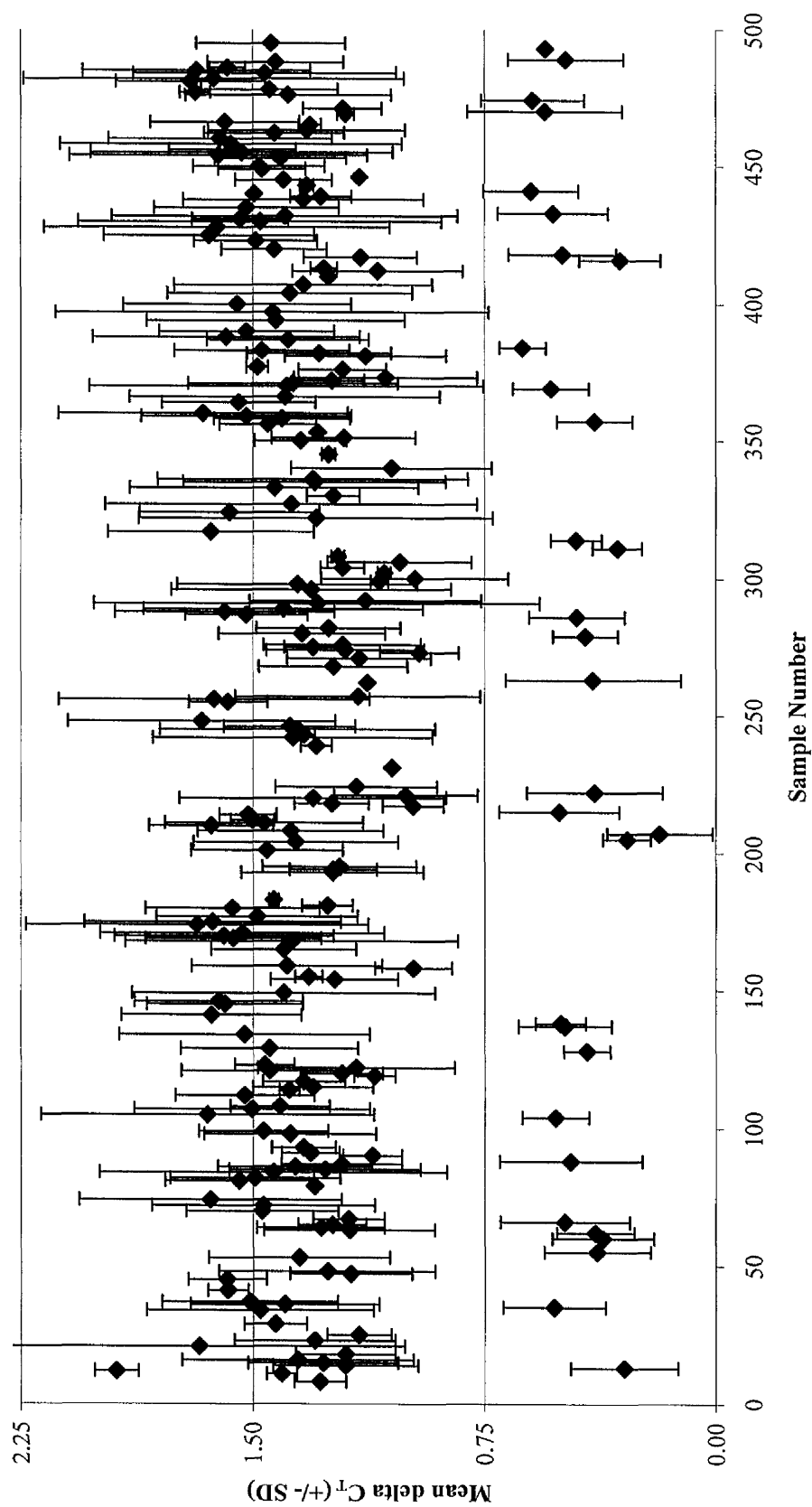
Figure 4.2 The mean ΔCT values ($\pm$ sd) for all subjects with the GSTM1 gene.

129

Table 4.5 GSTM1 gene copy number distribution in the study population.

GSTM1 Gene Copy Number	Cases	Controls					
		Bronchoscopy			Chest Clinic		
	N (%)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
0	94 (58.8)	104 (47.9)	1	1	39 (43.3)	1	1
1	57 (35.6)	99 (45.6)	0.64 (0.42-0.98)	0.63 (0.40-0.99)	43 (47.8)	0.55 (0.32-0.95)	0.56 (0.32-0.97)
2	9 (5.6)	14 (6.5)	0.71 (0.29-1.72)	0.67 (0.27-1.71)	8 (8.9)	0.47 (0.17-1.30)	0.43 (0.15-1.22)
1+2	66 (41.2)	113 (52.1)	0.65 (0.43-0.98)	0.64 (0.41-0.98)	51 (56.7)	0.54 (0.32-0.91)	0.54 (0.32-0.91)

[a]odds of lung cancer incidence (reference = GSTM1 null); [b]adjusted for age, sex and smoking duration.

trend test for bronchoscopy controls: $p = 0.07$

trend test for chest clinic controls: $p = 0.02$

4.2　GSTT1 Genotype Distribution

The GSTT1 genotype was determined in 99.4% of subjects; this represented 98.2% of cases (n = 161 / 164) and 100% of controls (n = 307 / 307). The cases not genotyped included one squamous cell carcinoma, one adenocarcinoma and one subject with no histology. In the three cases PCR products were not detectable despite repeated assays.

The GSTT1 null genotype was detected in 18.6% of cases (n = 30 / 161), 20.7% of bronchoscopy controls (n = 45 / 217) and 16.7% of chest clinic controls (n = 15 / 90). GSTT1 genotype was not associated with the odds of lung cancer incidence when cases were compared to bronchoscopy controls (OR_{adj} 0.90, 95% CI 0.53-1.55, p = 0.71) or chest clinic controls (OR_{adj} 1.18, 95% CI 0.59-2.36, p = 0.64) (Table 4.6). Subset analysis of GSTT1 distribution revealed no significant association with genotype and the odds of lung cancer when subjects were stratified by median age, gender or median smoking duration (Table 4.7). There was also no significant association of the GSTT1 genotype with lung cancer after cases were stratified according to histology (Table 4.8).

Table 4.6 GSTT1 gene distribution in the study population.

GSTT1	Cases N (%)	Controls Bronchoscopy N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	Controls Chest Clinic N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
Present	131 (81.4)	172 (79.3)	1	1	75 (83.3)	1	1
Null	30 (18.6)	45 (20.7)	0.88 (0.52-1.47)	0.90 (0.53-1.55)	15 (16.7)	1.15 (0.58-2.26)	1.18 (0.59-2.36)

[a]odds of lung cancer incidence (reference = GSTT1 present); [b]adjusted for age, sex and smoking duration

132

Table 4.7 GSTT1 gene distribution in the study population after stratification for: age, gender and smoking duration.

Factor	Strata	Cases		Controls					
				Bronchoscopy			Chest clinic		
		N +/- (%)		N +/- (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)	N +/- (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)
Age	≤ 65.5	39 / 14 (73.6/26.4)		88 / 21 (80.7/19.3)	1.50 (0.69-3.26)	1.47[cd] (0.67-3.24)	29 / 7 (80.6/19.4)	1.49 (0.53-4.15)	1.51[cd] (0.53-4.25)
	> 65.5	92 / 16 (85.2/14.8)		84 / 24 (77.8/22.2)	0.61 (0.30-1.22)	0.62[cd] (0.30-1.29)	46 / 8 (85.2/14.8)	1.00 (0.40-2.51)	0.93[cd] (0.37-2.36)
Gender	Male	88 / 21 (80.7/19.3)		107 / 23 (82.3/17.7)	1.11 (0.58-2.14)	1.17[bd] (0.59-2.30)	48 / 8 (85.7/14.3)	1.43 (0.59-3.48)	1.50[bd] (0.61-3.71)
	Female	43 / 9 (82.7/17.3)		65 / 22 (74.7/25.3)	0.62 (0.26-1.47)	0.57[bd] (0.22-1.43)	27 / 7 (79.4/20.6)	0.81 (0.27-2.42)	0.83[bd] (0.26-2.59)
Smoking Duration	≤ 40 yrs	35 / 9 (79.5/20.5)		93 / 23 (80.2/19.8)	1.04 (0.44-2.47)	1.02[bc] (0.43-2.43)	31 / 4 (88.6/11.4)	1.99 (0.56-7.12)	1.94[bc] (0.53-7.12)
	> 40 yrs	96 / 21 (82.1/17.9)		79 / 22 (78.2/21.8)	0.79 (0.40-1.53)	0.88[bc] (0.44-1.75)	44 / 11 (80.0/20.0)	0.88 (0.39-1.97)	0.98[bc] (0.42-2.26)

[a] odds of lung cancer incidence (reference = GSTT1 present); OR adjusted for [b] age, [c] sex and [d] smoking duration.

(GSTT1 present = (+), GSTT1 null = (-))

Table 4.8 GSTT1 gene distribution in the study population after stratification of cases according to histology.

Strata	Cases N + / - (%)	Controls Bronchoscopy N + / - (%)	Bronchoscopy Unadjusted[a] OR (95%CI)	Bronchoscopy Adjusted[ab] OR (95%CI)	Controls Chest clinic N + / - (%)	Chest clinic Unadjusted[a] OR (95%CI)	Chest clinic Adjusted[ab] OR (95%CI)
Squamous Cell	35 / 7 (83.3/16.7)		0.76 (0.32-1.83)	0.82 (0.33-2.06)		1.00 (0.37-2.67)	1.01 (0.36-2.83)
Adenocarcinoma	27 / 7 (79.4/20.6)	172 / 45 (79.3/20.7)	0.99 (0.41-2.42)	0.95 (0.38-2.34)	75 / 15 (83.3/16.7)	1.30 (0.48-3.52)	1.37 (0.49-3.79)
Small Cell	26 / 5 (83.9/16.1)		0.74 (0.27-2.02)	0.74 (0.26-2.11)		0.96 (0.32-2.91)	0.87 (0.27-2.75)
Large Cell	14 / 4 (77.8/22.2)		1.09 (0.34-3.48)	1.06 (0.33-3.41)		1.43 (0.41-4.95)	1.45 (0.41-5.10)
All NSCLC	76 / 18 (80.9/19.1)	172 / 45 (79.3/20.7)	0.91 (0.49-1.67)	0.92 (0.49-1.72)	75 / 15 (83.3/16.7)	1.18 (0.56-2.52)	1.20 (0.56-2.58)

[a] odds of lung cancer incidence (reference = GSTT1 present); [b] OR adjusted for age, sex and smoking duration.

134

4.3 hOGG1 Codon 326 Genotype Distribution

The hOGG1 codon 326 genotype was determined in 96.8% of the study population (n = 456 / 471). Subjects without genotype results included four cases, eight bronchoscopy and three chest clinic controls. The genotyping of these 15 subjects was not determined because of the failure of primary PCR. The distribution of hOGG1 codon 326 alleles did not differ from Hardy Weinberg equilibrium for both bronchoscopy (p = 0.88) and chest clinic controls (p = 0.64). The frequency of the rarer 326C allele was not significantly different between cases and either control group (Table 4.9).

Table 4.9 The distribution of the hOGG1 S326C genotype and frequency of the 326C allele within the study population.

Study Group	hOGG1 S326C genotype N (%)			Total	C allele	
	SS	SC	CC		frequency	p value
Cases	99 (61.9)	54 (33.8)	7 (4.4)	160	0.213	-
Bronchoscopy Controls	125 (59.8)	75 (35.9)	9 (4.3)	209	0.222	0.69
Chest Clinic Controls	57 (65.5)	24 (27.6)	6 (6.9)	87	0.207	0.57

Using the more common SS homozygotes as a reference, the odds of lung cancer incidence were calculated for SC heterozygotes and CC homozygotes (Table 4.10). Comparison with bronchoscopy controls revealed no significant association with the odds of lung cancer incidence for either heterozygotes (OR_{adj} 0.83, 95% CI 0.52-1.32; p = 0.44) or homozygotes (OR_{adj} 1.23, 95% CI 0.43-3.53; p = 0.71). This finding was consistent with case comparison to chest clinic controls, the adjusted odds ratio for heterozygotes was 1.29 (95% CI 0.72-2.32; p = 0.39) and homozygotes 0.66 (95% CI 0.21-2.11; p = 0.49). No significant association was found with lung cancer and carriers of the 326C allele (SC heterozygotes and CC homozygotes) compared to S326 homozygotes. No association was found with lung cancer for any genotype after stratification for gender, histology, median smoking duration and median age of controls (Table 4.10). A similar analysis was performed after cases were stratified according to histology, again no significant associations were found (Table 4.11).

135

Table 4.10 hOGG1 codon S326C polymorphism distribution in the study population.

hOGG1 Codon 326 Genotype	Cases N (%)	Controls				
		Bronchoscopy			Chest Clinic	
		Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
SS	99 (61.9)	1	1	57 (65.5)	1	1
SC	54 (33.8)	0.91 (0.59-1.41)	0.83 (0.52-1.32)	24 (27.6)	1.30 (0.73-2.32)	1.29 (0.72-2.32)
CC	7 (4.4)	0.98 (0.35-2.73)	1.23 (0.43-3.53)	6 (6.9)	0.67 (0.22-2.10)	0.66 (0.21-2.11)
SC + CC	61 (38.1)	1.09 (0.72-1.66)	1.15 (0.74-1.79)	30 (34.5)	0.85 (0.50-1.47)	0.86 (0.50-1.49)

[a]odds of lung cancer incidence (reference = SS homozygotes); [b]adjusted for age, sex and smoking duration

136

Table 4.11 hOGG1 codon S326C polymorphism distribution in the study population after stratification for: age, gender and smoking duration.

Factor	Strata	Cases	Controls					
			Bronchoscopy			Chest clinic		
		SS / SC+CC N (%)	SS / SC+CC N (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)	SS / SC+CC N (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)
Age	≤65.5	31 / 23 (57.4/42.6)	67 / 39 (63.2/36.8)	1.28 (0.65-2.49)	1.17[cd] (0.59-2.31)	18 / 16 (52.9/47.1)	0.84 (0.35-1.98)	0.74[cd] (0.30-1.82)
	>65.5	68 / 38 (64.2/35.8)	58 / 45 (56.3/43.7)	0.72 (0.41-1.26)	0.75[cd] (0.42-1.35)	39 / 14 (73.6/26.4)	1.56 (0.75-3.23)	1.58[cd] (0.75-3.31)
Gender	Male	70 / 38 (64.8/35.2)	78 / 48 (61.9/38.1)	0.88 (0.52-1.51)	0.83[bd] (0.48-1.44)	33 / 22 (60.0/40.0)	0.81 (0.42-1.59)	0.74[bd] (0.37-1.48)
	Female	29 / 23 (55.8/44.2)	47 / 36 (56.6/43.4)	1.04 (0.52-2.08)	0.95[bd] (0.44-2.03)	24 / 8 (75.0/25.0)	2.38 (0.90-6.27)	2.28[bd] (0.83-6.28)
Smoking Duration	≤40 yrs	28 / 17 (62.2/37.8)	69 / 41 (62.7/37.3)	1.02 (0.50-2.09)	0.97[bc] (0.47-2.02)	24 / 9 (72.7/27.3)	1.62 (0.61-4.29)	1.65[bc] (0.62-4.40)
	>40 yrs	71 / 44 (61.7/38.3)	56 / 43 (56.6/43.4)	0.81 (0.47-1.47)	0.84[bc] (0.48-1.47)	33 / 21 (61.1/38.9)	0.97 (0.50-1.89)	1.03[bc] (0.52-2.02)

[a] odds of lung cancer incidence (reference = SS homozygotes); OR adjusted for [b] age, [c] sex and [d] smoking duration.

Table 4.12 hOGG1 codon S326C polymorphism distribution in the study population after stratification of cases for histology.

Strata	Cases	Controls					
		Bronchoscopy			Chest clinic		
	SS / SC+CC N (%)	SS / SC+CC N (%)	Unadjusted[a] OR (95%CI)	Adjusted[ab] OR (95%CI)	SS / SC+CC N (%)	Unadjusted[a] OR (95%CI)	Adjusted[ab] OR (95%CI)
Squamous Cell	25 / 15 (62.5/37.5)		0.89 (0.45-1.79)	0.84 (0.40-1.76)		1.14 (0.52-2.48)	0.97 (0.42-2.23)
Adenocarcinoma	23 / 12 (65.7/34.3)	125 / 84 (59.8/40.2)	0.78 (0.37-1.65)	0.70 (0.32-1.50)	57 / 30 (65.5/34.5)	0.99 (0.43-2.27)	1.05 (0.45-2.42)
Small Cell	17 / 14 (54.8/45.2)		1.23 (0.57-2.62)	1.17 (0.53-2.57)		1.57 (0.68-3.60)	1.26 (0.52-3.07)
Large Cell	9 / 9 (50.0/50.0)		1.49 (0.57-3.90)	1.38 (0.52-3.66)		1.90 (0.68-5.29)	1.95 (0.69-5.53)
All NSCLC	57 / 36 (61.3/38.7)	125 / 84 (59.8/40.2)	0.94 (0.57-1.55)	0.86 (0.51-1.45)	57 / 30 (65.5/34.5)	1.20 (0.65-2.20)	1.17 (0.64-2.17)

[a]odds of lung cancer incidence (reference = SS homozygotes); [b]OR adjusted for age, sex and smoking duration.

4.4   MGMT Codon 178 Genotype Distribution

The MGMT codon 178 genotype was determined for 99.4% (n = 468 / 471) of the study population. One case and two bronchoscopy controls were not genotyped due to a failure of primary PCR; all chest clinic controls were genotyped. The distribution of codon K178R alleles in the case and control groups is shown in table 4.13.

Table 4.13 The distribution of codon 178 genotype within the study population.

| Study Group | | MGMT Codon 178 Genotype | | | Total | R allele | |
		KK	KR	RR		frequency	p value
Cases	n	124	38	1	163	0.123	-
	%	76.1	23.3	0.6			
Bronchoscopy Controls	n	161	50	4	215	0.135	0.79
	%	74.9	23.3	1.9			
Chest Clinic Controls	n	69	17	4	90	0.139	0.92
	%	76.7	18.9	4.4			

The frequency of the rarer 178R allele was 0.12 in the cases and 0.14 in both control groups, this difference was not significant. The distribution of alleles did not differ from Hardy Weinberg equilibrium for both bronchoscopy (p = 1.0) and chest clinic controls (p = 0.51). The odds of lung cancer incidence were calculated for heterozygotes (KR) and homozygotes (RR) for the 178R allele, by using KK homozygotes as a reference. Comparison with both the bronchoscopy (OR_{adj} 0.88, 95% CI 0.51-1.42; p = 0.61) and chest clinic controls (OR_{adj} 1.25, 95% CI 0.65-2.43; p = 0.50) revealed no significant association with odds of lung cancer incidence in KR heterozygotes (Table 4.14). The odds of lung cancer incidence in RR homozygotes were reduced, five fold when comparison was made with the bronchoscopy control group (OR_{adj} 0.19, 95% CI 0.02-1.91; p = 0.16) and over six fold when comparison was made with the chest clinic control group (OR_{adj} 0.15, 95% CI 0.02-1.46; p = 0.10) (Table 4.14). Neither finding reached significance, however of the nine RR homozygous subjects only one was a case.

There was no significant difference in genotype distribution between cases and chest clinic controls after stratification for age, gender and smoking duration (Table 4.15). However, the odds of lung cancer were reduced when cases were compared to bronchoscopy subjects with the 178R allele who had smoked for more than 40 years; this finding approached significance (OR_{adj} 0.57, 95% CI 0.31-1.06; p = 0.08).

Cases were stratified according to histology; a threefold reduction in odds of small cell lung cancer, that approached significance, was seen with subjects with the R allele when cases were compared to bronchoscopy controls (OR_{adj} 0.33, 95% CI 0.11-1.03; p = 0.06). No other significant associations were found (Table 4.16).

Table 4.14 MGMT codon 178 genotype distribution in the study population.

MGMT Codon 178 Genotype	Cases	Controls						
		Bronchoscopy			Chest Clinic			
	N (%)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
KK	124 (76.1)	161 (74.9)	1	1	69 (76.7)	1	1
KR	38 (23.3)	50 (23.3)	0.99 (0.61-1.60)	0.88 (0.51-1.42)	17 (18.9)	1.21 (0.64-2.37)	1.25 (0.65-2.43)
RR	1 (0.6)	4 (1.9)	0.33 (0.04-2.94)	0.19 (0.02-1.91)	4 (4.4)	0.14 (0.02-1.27)	0.15 (0.02-1.46)
KR + RR	39 (23.9)	54 (25.1)	0.94 (0.58-1.51)	0.82 (0.50-1.35)	21 (23.3)	1.03 (0.56-1.90)	1.04 (0.56-1.94)

[a]odds of lung cancer incidence (reference = KK homozygotes); [b]adjusted for age, sex and smoking duration.

141

Table 4.15 MGMT codon 178 genotype distribution in the study population after stratification for: age, gender and smoking duration.

Factor	Strata	Cases KK / KR+RR N (%)	Controls					
			Bronchoscopy			Chest clinic		
			KK / KR+RR N (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)	KK / KR+RR N (%)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)
Age	≤ 65.5	42 / 12 (77.8/22.2)	80 / 27 (74.8/25.2)	0.85 (0.39-1.84)	0.86[cd] (0.39-1.89)	27 / 9 (75.0/25.0)	0.86 (0.32-2.31)	0.84[cd] (0.31-2.28)
	> 65.5	82 / 27 (75.2/24.8)	81 / 27 (75.0/25.0)	0.99 (0.53-1.83)	1.11[cd] (0.62-1.98)	42 / 12 (77.8/22.2)	1.15 (0.53-2.50)	1.19[cd] (0.54-2.62)
Gender	Male	81 / 29 (73.6/26.4)	96 / 33 (74.4/25.6)	1.04 (0.58-1.86)	0.96[bd] (0.52-1.75)	46 / 10 (82.1/17.9)	1.65 (0.74-3.68)	1.80[bd] (0.79-4.13)
	Female	43 / 10 (81.1/18.9)	65 / 21 (75.6/24.4)	0.72 (0.31-1.68)	0.57[bd] (0.23-1.44)	23 / 11 (67.6/32.4)	0.49 (0.18-1.32)	0.41[bd] (0.15-1.17)
Smoking Duration	≤ 40 yrs	33 / 12 (73.3/26.7)	92 / 22 (80.7/19.3)	1.52 (0.68-3.41)	1.53[bc] (0.68-3.47)	24 / 11 (68.6/31.4)	0.79 (0.30-2.10)	0.78[bc] (0.29-2.07)
	> 40 yrs	91 / 27 (77.1/22.9)	69 / 32 (68.3/31.7)	0.64 (0.33-1.17)	0.57[bc] (0.31-1.06)	45 / 10 (81.8/18.2)	1.34 (0.60-3.00)	1.28[bc] (0.57-2.91)

[a]odds of lung cancer incidence (reference = KK homozygotes); OR adjusted for [b]age, [c]sex and [d]smoking duration.

142

Table 4.16 MGMT codon 178 genotype distribution in the study population after stratification of cases according to histology.

Strata	Cases KK / KR+RR N (%)	Controls Bronchoscopy KK / KR+RR N (%)	Bronchoscopy Unadjusted[a] OR (95%CI)	Bronchoscopy Adjusted[ab] OR (95%CI)	Controls Chest clinic KK / KR+RR N (%)	Chest clinic Unadjusted[a] OR (95%CI)	Chest clinic Adjusted[ab] OR (95%CI)
Squamous Cell	33 / 10 (76.7/23.3)		0.90 (0.42-1.96)	0.62 (0.27-1.43)		1.00 (0.42-2.35)	0.96 (0.38-2.44)
Adenocarcinoma	29 / 6 (82.9/17.1)	161 / 54 (74.9/25.1)	0.62 (0.24-1.57)	0.61 (0.24-1.56)	69 / 21 (76.7/23.3)	0.68 (0.25-1.86)	0.62 (0.22-1.72)
Small Cell	27 / 4 (87.1/12.9)		0.44 (0.15-1.32)	0.33 (0.11-1.03)		0.49 (0.15-1.55)	0.49 (0.14-1.65)
Large Cell	12 / 6 (66.7/33.3)		1.49 (0.53-4.16)	1.41 (0.50-4.01)		1.64 (0.55-4.91)	1.65 (0.54-5.04)
All NSCLC	74 / 22 (77.1/22.9)	161 / 54 (74.9/25.1)	0.89 (0.50-1.56)	0.77 (0.42-1.38)	69 / 21 (76.7/23.3)	0.98 (0.49-1.93)	0.98 (0.49-1.96)

[a] odds of lung cancer incidence (reference = KK homozygotes); [b] OR adjusted for age, sex and smoking duration.

143

4.5   MGMT Intron 1 (rs12268840) Genotype Distribution

The MGMT intron 1 (rs12268840) genotype was determined for 98.1% (n = 462 / 471) of the study population. A genotype result was not determined for two cases, four bronchoscopy and three chest clinic controls due to failure of primary PCR. The distribution and T allele frequency of the intron 1 genotype in case and control groups is shown in table 4.17. The distribution of alleles did not differ from Hardy Weinberg equilibrium for both bronchoscopy (p = 0.83) and chest clinic controls (p = 0.60). The frequency of the T allele was 0.29 in cases, 0.32 in bronchoscopy controls and 0.36 in chest clinic controls. The difference in allele frequency was significant between cases and chest clinic controls (p = 0.02) but not bronchoscopy controls (p = 0.20) (Table 4.17).

Table 4.17 The distribution of the intron 1 genotype within the study population.

Study Group		MGMT Intron 1 Genotype			Total	T allele	
		CC	CT	TT		frequency	p value
Cases	n	87	57	18	162	0.287	-
	%	53.7	35.2	11.1			
Bronchoscopy Controls	n	100	88	25	213	0.324	0.20
	%	46.9	41.3	11.7			
Chest Clinic Controls	n	33	46	8	87	0.356	**0.02**
	%	37.9	52.9	9.2			

The odds of lung cancer incidence were calculated for heterozygotes (CT) and homozygotes (TT) for the rarer intron 1 T allele, by using CC homozygotes as a reference. The distribution of the intron 1 genotype was not associated with a lung cancer incidence when cases were compared to bronchoscopy controls. The comparison of cases to chest clinic controls revealed a significant reduction in the odds of lung cancer incidence for heterozygous (CT) (OR_{adj} 0.47, 95% CI 0.26-0.82; p = 0.009) but not homozygous subjects (TT) (OR_{adj} 0.95, 95% CI 0.37-2.45; p = 0.92) (Table 4.18).

The analysis was repeated after stratification for age, gender and smoking duration, again no significant findings were seen with bronchoscopy controls (Table 4.19). Significant reductions in odds of lung cancer were seen with heterozygote subjects over 65.5 years (OR_{adj} 0.47, 95% CI 0.23-0.96), females (OR_{adj} 0.36, 95% CI 0.13-0.98) and those who had smoked for longer than 40 years (OR_{adj} 0.40, 95% CI 0.20-0.82) when cases were compared to chest clinic controls. These findings were not replicated in TT homozygotes (Table 4.20).

Cases were then stratified according to histology and the intron 1 genotype distribution re-analysed. There was a reduction in odds of squamous cell carcinoma of between three to five fold in case to bronchoscopy control heterozygote (OR_{adj} 0.29, 95% CI 0.12-0.68; p = 0.006) and homozygote comparison respectively (OR_{adj} 0.22, 95% CI 0.05-1.05; p = 0.06). This represented a significant trend for reducing odds of squamous cell carcinoma (p = 0.003) (Table 4.21). A significant reduction in odds of incidence of all NSCLC was also seen in heterozygotes (OR_{adj} 0.53, 95% CI 0.30-0.92) but this did not reach significance in homozygotes (OR_{adj} 0.47, 95% CI 0.18-1.20). This also represented a significant trend for reducing odds of NSCLC (p = 0.02).

Equivalent significant findings were seen in comparison to chest clinic control heterozygotes for both squamous cell carcinoma (OR_{adj} 0.18, 95% CI 0.07-0.48) and all NSCLC (OR_{adj} 0.35, 95% CI 0.18-0.66). A non-significant reduction in risk was seen for homozygotes and squamous cell carcinoma (OR_{adj} 0.30, 95% CI 0.06-1.59) and all NSCLC (OR_{adj} 0.52, 95% CI 0.17-1.56) (Table 4.22). There was a significant trend in the reduction of odds of both squamous cell carcinoma (p = 0.0007) and NSCLC (p = 0.008).

Table 4.18 MGMT intron 1 genotype distribution in the study population.

MGMT Intron 1 Genotype	Cases	Controls						
		Bronchoscopy				Chest Clinic		
	N (%)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)		N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
CC	87 (53.7)	100 (46.9)	1	1		33 (37.9)	1	1
CT	57 (35.2)	88 (41.3)	0.75 (0.49-1.17)	0.73 (0.46-1.16)		46 (52.9)	**0.47 (0.27-0.82)**	**0.47 (0.26-0.82)**
TT	18 (11.1)	25 (11.7)	0.84 (0.43-1.64)	0.89 (0.44-1.82)		8 (9.2)	0.85 (0.34-2.15)	0.95 (0.37-2.45)
CT + TT	75 (46.3)	113 (53.1)	0.76 (0.51-1.15)	0.76 (0.49-1.17)		54 (62.1)	**0.53 (0.31-0.90)**	**0.54 (0.32-0.93)**

[a] odds of lung cancer incidence (reference = CC homozygotes); [b] adjusted for age, sex and smoking duration.

Table 4.19 MGMT intron 1 genotype distribution in cases and bronchoscopy controls after stratification for: age, gender and smoking duration.

Factor	Strata	Cases CC / CT / TT N (%)	Bronchoscopy Controls					
			CC / CT / TT N (%)	CT		TT		
				Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)	Unadjusted[a] OR (95%CI)	Adjusted[a] OR (95%CI)	
Age	≤ 65.5	28 / 20 / 6 (51.9/37.0/11.1)	50 / 41 / 16 (46.7/38.3/15.0)	0.87 (0.43-1.77)	0.87[cd] (0.42-1.79)	0.67 (0.24-1.91)	0.64[cd] (0.22-1.86)	
	> 65.5	59 / 37 / 12 (54.6/34.3/11.1)	50 / 47 / 9 (47.2/44.3/8.5)	0.67 (0.38-1.18)	0.64[cd] (0.35-1.16)	1.13 (0.44-2.90)	1.28[cd] (0.47-3.51)	
Gender	Male	63 / 38 / 8 (57.8/34.9/7.3)	64 / 49 / 14 (50.4/38.6/11.0)	0.79 (0.46-1.36)	0.77[bd] (0.44-1.35)	0.58 (0.23-1.48)	0.62[bd] (0.24-1.63)	
	Female	24 / 19 / 10 (45.3/35.8/18.9)	36 / 39 / 11 (41.9/45.3/12.8)	0.73 (0.34-1.55)	0.67[bd] (0.30-1.50)	1.36 (0.50-3.71)	1.29[bd] (0.42-3.92)	
Smoking Duration	≤ 40 yrs	24 / 14 / 7 (53.3/31.1/15.6)	54 / 47 / 12 (47.8/41.6/10.6)	0.67 (0.31-1.44)	0.70[bc] (0.32-1.50)	1.31 (0.46-3.75)	1.37[bc] (0.48-3.95)	
	> 40 yrs	63 / 43 / 11 (53.8/36.8/9.4)	46 / 41 / 13 (46.0/41.0/13.0)	0.77 (0.43-1.36)	0.75[bc] (0.42-1.34)	0.62 (0.25-1.50)	0.61[bc] (0.24-1.53)	

[a] odds of lung cancer incidence (reference = CC homozygotes); [b] OR adjusted for age, sex and smoking duration.

147

Table 4.20 MGMT intron 1 distribution in cases and chest clinic controls after stratification for: age, gender and smoking duration.

Factor	Strata	Cases CC / CT / TT N (%)	Chest Clinic Controls					
			CC / CT / TT N (%)	CT Unadjusted[a] OR (95%CI)	CT Adjusted[a] OR (95%CI)	TT Unadjusted[a] OR (95%CI)	TT Adjusted[a] OR (95%CI)	
Age	≤65.5	28 / 20 / 6 (51.9/37.0/11.1)	12 / 18 / 4 (35.3/52.9/11.8)	0.48 (0.19-1.21)	0.46[cd] (0.18-1.19)	0.64 (0.15-2.70)	0.66[cd] (0.15-2.83)	
	>65.5	59 / 37 / 12 (54.6/34.3/11.1)	21 / 28 / 4 (39.6/52.8/7.5)	**0.47 (0.23-0.95)**	**0.47[cd] (0.23-0.96)**	1.07 (0.31-3.68)	1.35[cd] (0.37-4.90)	
Gender	Male	63 / 38 / 8 (57.8/34.9/7.3)	22 / 25 / 7 (40.7/46.3/13.0)	0.53 (0.26-1.07)	0.55[bd] (0.27-1.11)	0.40 (0.13-1.23)	0.44[bd] (0.14-1.35)	
	Female	24 / 19 / 10 (45.3/35.8/18.9)	11 / 21 / 1 (33.3/63.6/3.0)	0.42 (0.16-1.07)	**0.36[bd] (0.13-0.98)**	4.58 (0.52-40.4)	4.86[bd] (0.53-44.9)	
Smoking Duration	≤40 yrs	24 / 14 / 7 (53.3/31.1/15.6)	15 / 14 / 4 (45.5/42.4/12.1)	0.63 (0.23-1.67)	0.63[bc] (0.24-1.71)	1.09 (0.27-4.38)	1.13[bc] (0.28-4.59)	
	>40 yrs	63 / 43 / 11 (53.8/36.8/9.4)	18 / 32 / 4 (33.3/59.3/7.4)	**0.38 (0.19-0.77)**	**0.40[bc] (0.20-0.82)**	0.79 (0.22-2.77)	0.81[bc] (0.23-2.90)	

[a]odds of lung cancer incidence (reference = CC homozygotes); OR adjusted for [b]age, [c]sex and [d]smoking duration.

Table 4.21 MGMT intron 1 genotype distribution in cases stratified for histology and bronchoscopy controls.

Strata	Cases CC / CT / TT N (%)	Bronchoscopy Controls CC / CT / TT N (%)	CT Unadjusted[a] OR (95%CI)	CT Adjusted[ab] OR (95%CI)	TT Unadjusted[a] OR (95%CI)	TT Adjusted[ab] OR (95%CI)
Squamous Cell	31 / 9 / 2 (73.8/21.4/4.8)	100 / 88 / 25 (46.9/41.3/11.7)	**0.33 (0.15-0.73)**	**0.29 (0.12-0.68)**	0.26 (0.06-1.15)	0.22 (0.05-1.05)
Adenocarcinoma	18 / 14 / 2 (52.9/41.2/5.9)		0.88 (0.42-1.88)	0.84 (0.39-1.81)	0.44 (0.10-2.04)	0.44 (0.09-2.06)
Small Cell	14 / 14 / 3 (45.2/45.2/9.7)		1.14 (0.51-2.52)	1.09 (0.48-2.49)	0.86 (0.23-3.21)	0.77 (0.20-3.00)
Large Cell	9 / 6 / 3 (50.0/33.3/16.7)		0.76 (0.26-2.21)	0.71 (0.24-2.10)	1.33 (0.34-5.29)	1.32 (0.33-5.39)
All NSCLC	58 / 29 / 7 (61.7/30.9/7.4)	100 / 88 / 25 (46.9/41.3/11.7)	**0.57 (0.33-0.97)**	**0.53 (0.30-0.92)**	0.48 (0.20-1.19)	0.47 (0.18-1.20)

[a]odds of lung cancer incidence (reference = CC homozygotes); [b]OR adjusted for age, sex and smoking duration.

Trend test: squamous cell carcinoma (p = 0.003) and all NSCLC (p = 0.02).

Table 4.22 MGMT intron 1 genotype distribution in cases stratified for histology and chest clinic controls.

Strata	Cases CC / CT / TT N (%)	Chest Clinic Controls CC / CT / TT N (%)	CT Unadjusted[a] OR (95%CI)	CT Adjusted[ab] OR (95%CI)	TT Unadjusted[a] OR (95%CI)	TT Adjusted[ab] OR (95%CI)
Squamous Cell	31 / 9 / 2 (73.8/21.4/4.8)		**0.21 (0.09-0.50)**	**0.18 (0.07-0.48)**	0.27 (0.05-1.35)	0.30 (0.06-1.59)
Adenocarcinoma	18 / 14 / 2 (52.9/41.2/5.9)	33 / 46 / 8 (37.9/52.9/9.2)	0.56 (0.24-1.28)	0.56 (0.24-1.29)	0.46 (0.09-2.39)	0.52 (0.10-2.78)
Small Cell	14 / 14 / 3 (45.2/45.2/9.7)		0.72 (0.30-1.70)	0.67 (0.27-1.67)	0.88 (0.20-3.83)	1.04 (0.23-4.82)
Large Cell	9 / 6 / 3 (50.0/33.3/16.7)		0.48 (0.16-1.47)	0.47 (0.15-1.46)	1.38 (0.30-6.27)	1.40 (0.30-6.48)
All NSCLC	58 / 29 / 7 (61.7/30.9/7.4)	33 / 46 / 8 (37.9/52.9/9.2)	**0.36 (0.19-0.67)**	**0.35 (0.18-0.66)**	0.50 (0.17-1.50)	0.52 (0.17-1.56)

[a] odds of lung cancer incidence (reference = CC homozygotes); [b] OR adjusted for age, sex and smoking duration.

Trend test: squamous cell carcinoma (p = 0.0007) and all NSCLC (p = 0.008).

The combined MGMT genotype (codon 178 and intron 1) was analysed between cases and both control groups (Table 4.23). There was no difference in distribution between cases and bronchoscopy controls for any combination of MGMT genotype. The TT / AG genotype was not found in chest clinic controls, but was found in 5.0% of cases, this difference reached significance ($p = 0.05$). The CC / AA genotype was found less frequently and the CT / GG more frequently in chest clinic controls than cases, differences that approached significance ($p = 0.07$ and 0.09 respectively).

Table 4.23 Distribution of combined MGMT genotype (codon 178 and intron 1) distribution in cases and controls.

MGMT Genotype		Cases	Controls			
			Bronchoscopy		Chest Clinic	
Intron 1	Codon 178	N (%)	N (%)	p value	N (%)	p value
CC	AA	71 (44.1)	83 (39.2)	0.34	28 (32.2)	0.07
	AG	15 (9.3)	17 (8.0)	0.66	4 (4.6)	0.18
	GG	0	0	-	1 (1.1)	0.17
CT	AA	42 (26.1)	61 (28.8)	0.57	31 (35.6)	0.12
	AG	14 (8.7)	22 (10.4)	0.59	12 (13.8)	0.21
	GG	1 (0.6)	4 (1.9)	0.29	3 (3.4)	0.09
TT	AA	10 (6.2)	15 (7.1)	0.74	8 (9.2)	0.39
	AG	8 (5.0)	10 (4.7)	0.91	0	0.05[a]
	GG	0	0	-	0	-

[a]using Fishers exact test

4.6 Genotype Distribution after GSTM1 Stratification

The GSTM1 null genotype was a risk factor for lung cancer in this study (Section 4.1). The distribution of other genotypes was therefore assessed in the study population after stratification for the presence or absence of the GSTM1 gene (Table 4.24). GSTM1 stratification had a significant impact on the distribution of GSTT1 genotype according to case status, but no significant impact on other genotypes.

When the GSTM1 gene was present, individuals null for GSTT1 were at increased odds of lung cancer compared to those with GSTT1. This reached significance when cases were compared to bronchoscopy controls (OR_{adj} 2.55, 95% CI 1.10-5.88, p = 0.03) but not chest clinic controls (OR_{adj} 2.10, 95% CI 0.77-5.72, p = 0.15). The reverse was true when the GSTM1 gene was absent, individuals null for GSTT1 were at decreased odds of lung cancer compared to those with GSTT1. This was a significant finding when cases were compared to bronchoscopy controls (OR_{adj} 0.39, 95% CI 0.19-0.83, p = 0.01) but not with chest clinic controls (OR_{adj} 0.69, 95% CI 0.26-1.84, p = 0.46).

The interaction of GSTM1, GSTT1 and the odds of lung cancer was analysed by categorising subjects into the four possible combinations of genotype (M+/T+, M+/T-, M-/T+ and M-/T-) and using the M+/T+ category as a reference (Table 4.25). Subjects with either the M+/T- or M-/T+ genotype combination had a greater than twofold increased odds of lung cancer incidence. The adjusted odds ratio for the M+/T- combination was 2.56 (95% CI 1.11-5.91, p = 0.03) when cases were compared to bronchoscopy controls and 2.24 (95% CI 0.82-6.07, p = 0.11) when comparison was made to chest clinic controls. Similar odds were seen for subjects with the M-/T1+ genotype, the adjusted odds ratio for the bronchoscopy control comparison was 2.29 (95% CI 1.40-3.75, p = 0.001) and 2.32 (95% CI 1.29-4.18, p = 0.005) after comparison was made with chest clinic controls. The homozygous absence of both genes was not associated with a significantly different incidence of lung cancer; the adjusted odds ratio of lung cancer was 0.91 (95% CI 0.43-1.92, p = 0.80) when cases were compared to bronchoscopy controls. Although the adjusted odds of lung cancer were increased in the chest clinic control comparison, this finding did not reach significance (OR_{adj} 1.56, 95% CI 0.59-4.14, p = 0.37).

Table 4.24 Adjusted odds of lung cancer by genotype after study population stratification according to GSTM1 status.

Genotype / Phenotype	GSTM1			
	Present		Null	
	Bronchoscopy controls	Chest clinic controls	Bronchoscopy controls	Chest clinic controls
	OR$_{adj}$[a] (95% CI)	OR$_{adj}$[a] (95% CI)	OR$_{adj}$[a] (95% CI)	OR$_{adj}$[a] (95% CI)
GSTT1 null	**2.55 (1.10-5.88)**	2.10 (0.77-5.72)	**0.39 (0.19-0.83)**	0.69 (0.26-1.84)
hOGG1 codon 326 (SC)	0.60 (0.29-1.24)	0.98 (0.41-2.35)	1.02 (0.54-1.92)	1.53 (0.66-3.56)
hOGG1 codon 326 (CC)	1.20 (0.18-7.78)	0.47 (0.07-3.02)	1.23 (0.34-4.52)	0.90 (0.18-4.41)
MGMT codon 178 (KR)	0.92 (0.44-1.92)	2.19 (0.81-5.90)	0.91 (0.43-1.95)	0.66 (0.26-1.67)
MGMT codon 178 (RR)	-	-	0.24 (0.02-2.98)	0.28 (0.02-3.52)
MGMT intron 1 (CT)	0.73 (0.37-1.47)	0.44 (0.19-1.01)	0.69 (0.36-1.32)	**0.42 (0.18-0.97)**
MGMT intron 1 (TT)	0.77 (0.26-2.25)	0.74 (0.20-2.69)	1.08 (0.40-2.91)	1.02 (0.23-4.52)

[a] odds of lung cancer incidence adjusted for age, gender and smoking duration

153

Table 4.25 Distribution of the combined GSTM1 and GSTT1 genotype in the study population.

GSTM1 and GSTT1 Genotype	Cases	Controls					
		Bronchoscopy			Chest Clinic		
	N (%)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)
M1 + / T1 +	50 (31.3)	99 (45.6)	1	1	44 (48.9)	1	1
M1 + / T1 -	16 (10.0)	14 (6.5)	2.26 (1.02-5.01)	2.56 (1.11-5.91)	7 (7.8)	2.01 (0.76-5.34)	2.24 (0.82-6.07)
M1 - / T1 +	80 (50.0)	73 (33.6)	2.17 (1.36-3.46)	2.29 (1.40-3.75)	31 (34.4)	2.27 (1.27-4.06)	2.32 (1.29-4.18)
M1 - / T1 -	14 (8.8)	31 (14.3)	0.89 (0.44-1.83)	0.91 (0.43-1.92)	8 (8.9)	1.54 (0.59-4.02)	1.56 (0.59-4.14)

[a]odds of lung cancer incidence (reference = M1+/T1+); [b]adjusted for age, sex and smoking duration

(+) = gene present, (-) = null

154

The interaction of GSTM1 gene copy number and GSTT1 genotype was examined by stratifying the study population according to GSTM1 copy number. The odds of lung cancer in individuals null for GSTT1 were calculated in each stratum relative to those with GSTT1. The comparison of cases to bronchoscopy controls revealed an increased adjusted odds of lung cancer incidence of 6.62 (95% CI 0.37-118.2; p = 0.20) in GSTT1 null subjects with two copies of GSTM1, 2.13 (95% CI 0.83-5.48; p = 0.12) in GSTT1 null subjects with one copy of GSTM1 and 0.39 (95% CI 0.19-0.83; p = 0.01) in GSTT1 and GSTM1 null subjects (Table 4.26); this was a significant trend (p = 0.005). Cases were also compared to chest clinic controls, no odds ratio was calculable in the two GSTM1 copy subset, the adjusted odds in the one copy subset was 1.32 (95% CI 0.45-3.89; p = 0.61) and for subjects null for both genes was 0.69 (95% CI 0.26-1.84; p = 0.46) (Table 4.26). Analysis of the bronchoscopy control comparison suggested that the odds of lung cancer increased with the presence of two copies of GSTM1 compared to one copy in subjects null for GSTT1. However the wide confidence intervals are an indication that although an effect is possible the study is underpowered to determine the effect with accuracy.

Table 4.26 Distribution of the GSTT1 genotype in the study population after stratification according to GSTM1 copy number.

GSTM1 Copy Number	Cases	Controls						
		Bronchoscopy			Chest Clinic			
	GSTT1 +/- N (%)	GSTT1 +/- N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	GSTT1 +/- N (%)	Unadjusted[a] OR (95% CI)	Adjusted[ab] OR (95% CI)	
0	31 / 8 (85.1/14.9)	73 / 31 (70.2/29.8)	**0.41 (0.20-0.84)**	**0.39 (0.19-0.83)**	31 / 8 (79.5/20.5)	0.68 (0.26-1.78)	0.69 (0.26-1.84)	
1	46 / 11 (80.7/19.3)	88 / 11 (88.9/11.1)	1.91 (0.77-4.75)	2.13 (0.83-5.48)	36 / 7 (83.7/16.3)	1.23 (0.43-3.49)	1.32 (0.45-3.89)	
2	4 / 5 (44.4/55.6)	11 / 3 (78.6/21.4)	4.58 (0.73-28.65)	6.62 (0.37-118.19)	8 / 0 (100/0)	-	-[c]	

[a] odds of lung cancer incidence (reference = GSTT1 present); [b] adjusted for age, sex and smoking duration

[c] using Fishers exact test: p = 0.03

trend test for bronchoscopy control p = 0.005

4.7 Discussion

This study was designed to investigate whether genetic variability might influence susceptibility to lung cancer development. The homozygous null GSTM1 genotype was shown to increase the risk of developing lung cancer. This was a consistent finding when cases were compared to both referent groups and also consistent with three previous meta-analyses all of which reported a weak association with lung cancer (Houlston 1999; Benhamou, Lee et al. 2002; Ye, Song et al. 2006). The effect of the null polymorphism was marginally stronger in this study than that reported by the meta-analyses (e.g. OR 1.18 95% CI 1.14-1.23 reported by Song *et al* or 1.10 95% CI 1.01-1.19 by Benhamou *et al*); one reason for this may have been the high level of tobacco exposure reported by study participants as a whole. For example, in a recent study by Sorensen *et al*, no association between the GSTM1 null genotype and lung cancer was found. However, the proportion of never smokers was significantly greater in controls (34%) compared to cases (4%); the length and intensity of smoking was also significantly greater in cases than controls (Sorensen, Raaschou-Nielsen et al. 2007). Studies using such controls with a lower tobacco exposure might dilute the protective effect of studied genotypes by comparing cases to individuals who would have developed lung cancer given a comparable tobacco history. A control group which has been heavily exposed to tobacco and not developed lung cancer, is in essence relatively resistant to tobacco's carcinogenic properties. The genotype distribution of this population would provide a greater contrast to cases, by definition relatively tobacco sensitive, than a non-exposed healthy population (Belogubova, Togo et al. 2004). Ideally a referent group should have a similar exposure history to cases; it could be postulated that the reason for controls not developing lung cancer could therefore in theory be genetic and not environmental.

A general observation through all subset analyses was that the GSTM1 null genotype adversely affected the risk of lung cancer. The results of histological subset analysis showed significantly increased odds of lung cancer when all cases with NSCLC were compared to bronchoscopy (OR_{adj} 1.70, 95% CI 1.02-2.85, p = 0.04) and chest clinic controls (OR_{adj} 2.05, 95% CI 1.12-3.73, p = 0.02). The specific histological type most strongly associated with the null genotype was large cell carcinoma, which had adjusted odds ratios of 4.80 (95% CI 1.44-16.0, p = 0.01) and 3.99 (95% CI 1.26-12.6, p = 0.02) when compared to bronchoscopy and chest clinic controls respectively. The wide confidence intervals seen are an indication of the small numbers involved and therefore

caution in interpretation of size of effect is required. Exposure to polycyclic aromatic hydrocarbons (PAH) has previously been linked to the development of squamous cell carcinoma (Hoffmann, Djordjevic et al. 1997); a lack of GSTM1 which detoxifies the activated form of benzo[a]pyrene, has also been associated with squamous cell carcinoma (Belogubova, Ulibina et al. 2006). The association of the null genotype with large cell carcinoma is a novel finding and might suggest a similar association with the carcinogenic properties of PAH as squamous cell carcinoma.

There was evidence for the null genotype having a greater effect in subjects over the median age in the chest clinic arm (OR_{adj} 2.45, 95% CI 1.24-4.84, p = 0.01) and to a lesser degree in bronchoscopy controls (OR_{adj} 1.62, 95% CI 0.92-2.84, p = 0.09). The increased effect seen in older individuals may reflect the relative resistance of older chest clinic controls to tobacco carcinogens compared to cases.

One aim of this study was to address the issue of GSTM1 copy number and lung cancer risk. Subjects with two copies of GSTM1 have been shown to have significantly greater enzymic activity than subjects with one copy (McLellan, Oscarson et al. 1997). The pre-study hypothesis was therefore that subjects with two copies would be less likely to develop the disease and so demonstrate a protective dose effect of the GSTM1 gene. The only previous study examining lung cancer and GSTM1 copy number found no association between the presence or absence of GSTM1 and lung cancer and therefore no gene dose effect was detected (Sorensen, Raaschou-Nielsen et al. 2007). However, the control group in this study, as previously discussed, had significantly less tobacco exposure than cases which may have resulted in an underestimation of the effect of the GSTM1 null genotype.

The determination of GSTM1 gene copy number was achieved by multiplex real time PCR, using relative quantitation of the GSTM1 gene to a housekeeping gene (albumin) following a method detailed by Brasch-Andersen *et al* (Brasch-Andersen, Christiansen et al. 2004). Using three standard deviations of the mean difference between GSTM1 and albumin, it was possible to define two mathematically distinct groups. The proportion of controls with two copies (7.2%, n = 22 / 307), one copy (46.3%, n = 142 / 307) and no copies (46.6%, n = 143 / 307) of GSTM1 was similar to a previous study investigating genetic susceptibility to asthma (Brasch-Andersen, Christiansen et al. 2004) and a study investigating GSTM1 and breast cancer risk (Roodi, Dupont et al. 2004). The finding of

158

three individuals (1 case and 2 controls) with very high ΔC_T was unexpected. Previous work by McLellan *et al*, had identified two individuals with ultra-rapid GSTM1 enzyme activity, these individuals had three copies of the GSTM1 gene (McLellan, Oscarson et al. 1997). However, this would have resulted in negative ΔC_T values. Increased copy number of the albumin gene may have resulted in increased ΔC_T values, the use of a different housekeeping gene could be one way of further testing this possibility. The finding remains unexplained at present.

The odds of lung cancer incidence were determined using GSTM1 null subjects as a reference. There was evidence of a significant dose effect when cases were compared to chest clinic controls (trend test $p = 0.02$); this effect approached significance when cases were compared to bronchoscopy controls (trend test $p = 0.07$). The chest clinic comparison revealed a lower risk of lung cancer in subjects with two copies of the GSTM1 gene compared to subjects with one copy and no copies. The lower effect in bronchoscopy controls may be due to lower tobacco exposure in this group. However, due to the low frequency of GSTM1 homozygotes a larger study would be required to confirm these findings.

The distribution of the GSTT1 null genotype, a homozygous null deletion polymorphism present in approximately 20% of Caucasians, was investigated in the study population (Garte, Gaspari et al. 2001). The null genotype was detected in 18.6% ($n = 30 / 161$) of cases, 20.7% ($n = 45 / 217$) of bronchoscopy controls and 16.7% ($n = 15 / 75$) of chest clinic controls. There was no significant association with lung cancer overall or after subset analysis. This was consistent with a recent meta-analysis and also with a previous study in the Wythenshawe population (Lewis, Cherry et al. 2002; Raimondi, Paracchini et al. 2006). A positive association reported by a second recent meta-analysis included studies from ethnically diverse populations (Ye, Song et al. 2006); the frequency of the GSTT1 null genotype is highly variable between different ethnic groups and may explain the difference in findings between the two meta-analyses (Nelson, Wiencke et al. 1995).

There was however evidence for a significant interaction with the GSTM1 genotype, especially in reference to bronchoscopy controls. The presence of either the GSTM1 or GSTT1 gene in the absence of the other doubled the odds of lung cancer incidence when compared to subjects with both genes. This contrasted with no significant increase in odds of lung cancer incidence in subjects null for both. This was a surprising finding, the pre-

study hypothesis was that subjects null for both genes might have an increased likelihood of lung cancer. Although a similar interaction has previously been reported by Ruano-Ravina *et al* (Ruano-Ravina, Figueiras et al. 2003), other studies have reported no interaction (Nazar-Stewart, Vaughan et al. 2003), or increased risk of lung cancer with the presence of both GSTM1 and GSTT1 null polymorphisms (Kiyohara, Yamamura et al. 2000; Sorensen, Autrup et al. 2004). To further analyse the relationship between GST genotype and lung cancer, the study population was stratified according to GSTM1 copy number and the risk of lung cancer calculated using GSTT1 present subjects as a reference. A significant trend was seen with case comparison to bronchoscopy controls (p = 0.005). There was a dose effect with increasing GSTM1 copy number and risk of lung cancer in GSTT1 null subjects.

Both genes can facilitate the activation and detoxification of carcinogens (Hayes, Flanagan et al. 2005) and also the removal of protective chemicals such as isothiocyanates (Brennan, Hsu et al. 2005). A study by London *et al*, demonstrated a protective effect of increased isothiocyanate levels against lung cancer risk especially in individuals null for both GSTT1 and GSTM1 (London, Yuan et al. 2000); these findings were supported by a recent study by Brennan *et al* (Brennan, Hsu et al. 2005). Subjects in a further study, with reduced reported isothiocyanate intake were at increased risk of lung cancer with the double null genotype (Spitz, Duphorne et al. 2000). Variations in diet between different study populations may therefore explain some of the inconsistency in results seen. Variation in reported occupational exposure may be another factor. A recent meta-analysis of the GSTT1 null polymorphism showed no overall association with lung cancer. However an increased risk was found in individuals who reported occupational exposure especially to asbestos (Raimondi, Paracchini et al. 2006). Cases and both control groups from the current study reported a high degree of occupational exposure which may have been a factor in influencing the interaction between GSTT1 and GSTM1.

The GSTM1 null genotype has been associated with an increased risk of severe COPD (Cheng, Yu et al. 2004), the GSTT1 and GSTM1 null genotypes have in some studies been associated with increased risk of asthma (Ivaschenko, Sideleva et al. 2002; Brasch-Andersen, Christiansen et al. 2004) and also in combination an increased risk of COPD (Calikoglu, Tamer et al. 2006). Bronchoscopy controls were significantly more likely to have a diagnosis of asthma than cases (12.0% vs. 6.1%, p = 0.05) and the most common diagnosis of chest clinic controls was COPD (80.0% vs. 27.4%, p = 0.0009). The presence

of COPD and asthma in controls may have resulted in selection of individuals more likely to have a GST null genotype than a healthy population. The impact of such a bias would have been an underestimation of the effect of GST null genotypes on lung cancer risk.

There was no significant difference in the distribution of the hOGG1 polymorphism Ser326Cys between cases and either control group, in the whole study population, after subset analysis or after stratification by GSTM1 status. Reduced hOGG1 repair activity, in lymphocytes, has been associated with an increased risk of lung cancer (Gackowski, Speina et al. 2003; Paz-Elizur, Krupsky et al. 2003). A polymorphism associated with reduced hOGG1 repair activity might therefore also be associated with increased lung cancer risk. A recent review of studies investigating the functional impact of the Ser326Cys polymorphism on hOGG1 activity showed that out of fourteen studies examined only three reported a positive impact of the polymorphism on function; with the Cys polymorphism resulting in a decreased ability to repair 8-oxoG (Weiss, Goode et al. 2005). Inconsistent findings from functional investigations have been mirrored in results from epidemiological studies (reviewed in Section 1.4.2.1). The largest study of the hOGG1 codon 326 polymorphism, which examined the genotype distribution of 2188 cases and 2198 controls, did not show an increased risk of lung cancer for Cys/Cys homozygotes compared to Ser/Ser homozygotes (Hung, Brennan et al. 2005). Although, an increased risk of adenocarcinoma was reported. The lack of association between the hOGG1 codon 326 polymorphism and lung cancer is therefore consistent with reports that the polymorphism has no functional effect on protein function. It is also consistent with mice models that report no overt phenotype with OGG$^{-/-}$ null mice (Klungland, Rosewell et al. 1999; Osterod, Hollenbach et al. 2001; Arai, Kelly et al. 2006).

The two single nucleotide polymorphisms examined in the MGMT gene were chosen because previous work by Margison *et al* had shown that these sites were most significantly associated with MGMT activity (Margison, Heighway et al. 2005). The codon 178 polymorphism involved a lysine (AAG) to arginine (AGG) substitution, which is in 100% linkage equilibrium with an isoleucine (ATC) to valine (GTC) substitution in codon 143. This is thought to be important because of its close proximity to the central cysteine acceptor residue at codon 145. Functional assessment of the codon 178 variants have suggested that the lysine (K178) variant is more easily inactivated than the arginine (178R) variant, but that both repair O^6-meG with similar efficacy (Margison, Heighway et al. 2005). A functional difference has also recently been reported between the K178 and

178R variant in the repair of O^6-[4-oxo-4-(3-pyridyl)butyl]guanine (O^6-pobG), a pyridyloxobutyl adduct formed by activated NNK (Mijal, Kanugula et al. 2006). The repair of this adduct was reported to be profoundly effected by sequence context, although the 178R protein was less sensitive to sequence specificity than the K178 protein (Mijal, Kanugula et al. 2006). The pre-study hypothesis was that subjects with the 178R variant would be less susceptible to alkylating damage and would therefore be less likely to develop lung cancer.

The main finding from the study of the K178R polymorphism was that there was no significant association with lung cancer incidence in the whole study population, after subset analysis and after stratification by GSTM1 status. However, subjects homozygous for the 178R allele had a marked reduction in odds of lung cancer incidence, in both chest clinic (OR_{adj} 0.15, 95% CI 0.02-1.46, p = 0.10) and bronchoscopy control (OR_{adj} 0.19, 95% CI 0.02-1.91, p = 0.16) comparisons compared to subjects homozygous for the K178 variant. This observation did not reach significance because of the rarity of RR homozygotes, but it was interesting to note that of nine RR homozygotes in the study only one was a case. The lack of reduced risk in heterozygotes might suggest that the effect of the K178 variant may be dominant over the protective effect of the 178R variant in this study population. It was interesting to note that when allelic expression imbalance was measured in blood using the MGMT codon 178 polymorphism as a flag, 87.5% of samples where imbalance was detected were directed towards relative overexpression of the K178 variant protein (Section 7.1), a similar finding was reported previously by Margison *et al* (Margison, Heighway et al. 2005).

Exon 5 MGMT polymorphisms (codon 143/178) and lung cancer risk have been investigated in four previous case-control studies (Kaur, Travaline et al. 2000; Cohet, Borel et al. 2004; Krzesniak, Butkiewicz et al. 2004; Yang, Coles et al. 2004). The results from these studies have been inconclusive (Table 4.27). Two studies have shown an increase in risk, one in non-smokers (Cohet, Borel et al. 2004) and the second in the whole study population but not in the Caucasian only subpopulation (Kaur, Travaline et al. 2000). Two further studies have shown no association (Krzesniak, Butkiewicz et al. 2004; Yang, Coles et al. 2004). All of these studies were small and therefore lacked power to conclusively examine MGMT polymorphisms and lung cancer risk. The reported allele frequency for the 178R allele in Caucasians ranged from 0.11 to 0.13 in cases and 0.07 to 0.12 in controls; this compares to 0.12 in cases and 0.14 in both control groups in this

study. The allele frequency is therefore higher in both control groups for this study than previous studies, which may reflect the higher smoking prevalence and age of the controls.

Table 4.27 Summary of studies examining exon 5 SNPs of MGMT and lung cancer.

Author/ Country	Cases / Controls	Smokers (%) Case vs Control	Allele Frequency Case vs Control	OR (KR)	Comments
Kaur 2000 USA	139 / 139 (56 Caucasian) Hospital	89 vs 71	0.11 vs 0.07	2.1 (2.0 Cauc)	Increased risk – but not in Caucasians only
Cohet 2004 8 countries	136 / 133 Caucasian Hosp / Pop	0	0.13 vs 0.07	2.05	All non-smokers Increased risk – esp adenoca.
Krzesniak 2004 Poland	96 / 96 Caucasian Population	82	0.12 vs 0.12	1.0	No association
Yang 2004 USA	92 / 85 Caucasian Hospital	74	0.11 vs 0.09	1.19	No association
Current Study UK	164 / 307 Caucasian Hospital	96 vs 92	0.12 vs 0.14	RR 0.19+0.15	Decreased risk for RR – not significant

The effect of the K178R polymorphism has also been recently investigated in a nested case-control study of endometrial cancer (456 cases and 1134 controls, all female). The 178R allele frequency in this study was 0.12 in both cases and controls. Overall there was no association with K178R and endometrial cancer risk. However, KR heterozygotes with a greater than 30 packyear history were at reduced risk of endometrial cancer (Han, Hankinson et al. 2006). A similar nested case-control study showed a protective effect of the 178R allele (allele frequency 0.07 in cases and 0.14 in controls) and colorectal cancer in females but not males (Tranah, Bugni et al. 2006) and a further study showed no association with breast cancer risk (Han, Tranah et al. 2006).

Although the current study of codon 178 polymorphisms and lung cancer is the largest undertaken, a much larger study would be required to determine the true impact of the RR genotype on lung cancer incidence. However, the results are suggestive of a strong protective effect against lung cancer development and therefore further study in this area is warranted.

The second single nucleotide polymorphism of the MGMT gene investigated in this study involved a single base change in the first intron (C to T), a non-coding part of the gene. No phenotypic change in enzyme is thought to result directly from this change however this position serves as a marker for an effect on MGMT activity possible through expression levels (Margison, Heighway et al. 2005). The T allele was found significantly more frequently in the chest clinic control population than in cases (0.36 vs. 0.29, p = 0.02), the frequency of the T allele in bronchoscopy controls was greater than cases but this difference did not reach significance (0.32 vs. 0.29, p = 0.20). Heterozygote subjects (CT) were found to be significantly less likely to have lung cancer when cases were compared to chest clinic controls (OR_{adj} 0.47, 95% CI 0.26-0.82, p = 0.009) this finding was not replicated in bronchoscopy controls. There was no significant association with lung cancer for homozygotes with the rarer allele (TT) in either study group.

Histological subset analysis showed a significant trend for reduced risk of squamous cell carcinoma in subjects heterozygous and homozygous for the T allele when cases were compared to bronchoscopy controls (p = 0.003) and chest clinic controls (p = 0.0007). The highly significant reduction in risk of squamous cell carcinoma is a novel finding and suggests a possible role of MGMT in the development of this histological type of lung cancer. Silencing of the MGMT gene has been reported to be a more frequent occurrence in lung adenocarcinomas than squamous cell carcinomas (Wolf, Hu et al. 2001). Exposure to the carcinogenic methylating agent NNK is also thought to be a potent inducer of adenocarcinoma (Hecht 1998). The reason for an association between the intron 1 polymorphism and squamous cell carcinoma is therefore unclear. Caution should be used in the interpretation of a result from subset analysis, especially with the relatively small number of squamous cell cases analysed. A larger study would be required to confirm these findings.

Chapter Five

The Development and Validation of MGMT and MPG Oligonucleotide Cleavage Assays

5.1 MGMT and MPG Assay Development

5.1.1 Introduction

The measurement of MGMT and MPG repair activity, using two separate [^{32}P]-labelled oligonucleotide cleavage assays, had been performed in the Carcinogenesis Group of the Paterson Institute for Cancer Research previously but not for processing large numbers of samples. The principle of both assays was the measurement of the amount of [^{32}P]-labelled fragment cleaved from a double stranded oligonucleotide, which was fixed to a microtitre plate well via a biotin-streptavidin bond, through the repair of specific DNA adducts (Figure 5.1).

In the case of MGMT, this was an indirect measurement. The assay, which was based on work by Wu *et al* (Wu, Hurst-Calderone et al. 1987), involved two steps. During the first step, the sample extract was incubated with substrate oligonucleotide containing O^6-methylguanine (O^6-meG). MGMT operates by the stoichiometric removal of the methyl group from O^6-meG leaving a normal guanine base *in situ* and inactivating the DNA repair protein. The repair of O^6-meG in this assay created a site for digestion by the restriction enzyme *Pst*1. The second step involved the washing of the microtitre plate to remove extract before the introduction of *Pst*1. The activity of MGMT was calculated by the amount of radiolabelled 5mer released by *Pst*1. This was directly related to the number of *Pst*1 sites created which was related to the number of O^6-meG adducts repaired. As MGMT is inactivated by the repair of a single adduct, the assay was essentially measuring the number of active MGMT molecules in an extract.

The MPG assay was modified from the method used by Elder *et al* (Elder, Jansen et al. 1998). The repair of an ethenoadenine (εA) adduct in the presence of activated APE-1 resulted in the cleavage of a [^{32}P]-labelled 5mer from the double stranded oligonucleotide. The amount of radiolabelled 5mer released was used as a measure of MPG activity. MPG is a true enzyme and therefore its activity was calculated as a measure of adducts repaired over time. The method for both assays required development and validation.

(A)
Common factors to both assays include: substrate oligonucleotide containing a specific DNA adduct (εA for MPG and O^6-meG for MGMT) hybridised to a complement strand which is attached to the reaction well via a strong streptavidin-biotin bond.
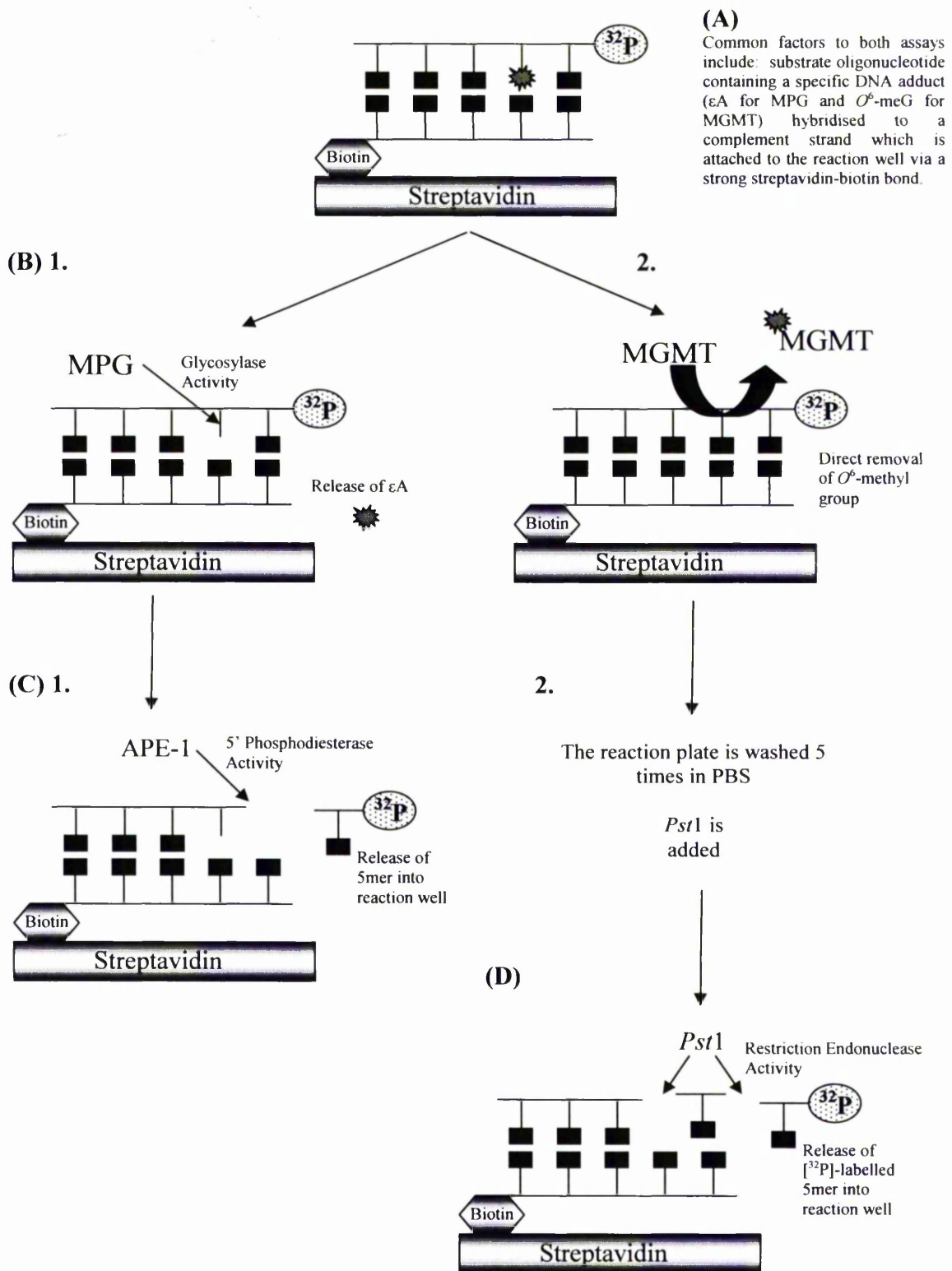
Figure 5.1 A diagrammatic representation of the MGMT and MPG oligonucleotide cleavage activity assays.

(A) see above;

(B) 1. MPG glycosylase activity hydrolyses the glycosidic bond and so removes εA (leaving an AP site).

2. MGMT directly removes the O^6-methyl group leaving a normal guanine residue and inactivating the repair protein.

(C) 1. Activated APE-1 cleaves the phosphodiester bond immediately 5' to an AP site and so releases ^{32}P into the reaction well.

2. The MGMT reaction plate is washed prior to the addition of the restriction enzyme *Pst*1.

(D) In the presence of a normal guanine a restriction site allows cleavage by *Pst*1 releasing a [^{32}P]-labelled 5mer into the reaction well, if an O^6-meG residue persists cleavage is prevented and no radioactivity is released.

167

5.1.2   Screening of Cell Lines

To aid in the process of method development and generate a quality control sample for study assays, cell lines were screened for MGMT and MPG activity. MCF-7 cells were chosen since they expressed measurable levels of both enzymes. MCF-7 cells were grown, by A Watson, as a monolayer in RPMI containing 10% FCS and 2mM L-Glutamine and incubated at 37°C in 5% CO_2 / 95% air. Cells were harvested, pooled, washed with PBS, aliquoted (~5 million cells / aliquot), centrifuged and the supernatant removed before pellets were stored at -20°C.

5.1.3   Substrate Binding

Initial assays showed marked variation of negative control values i.e. the amount of radioactivity released by wells when incubated with reaction buffer only. There was also a lack of linearity in values when serial dilutions of individual samples were assayed. The radioactive labelling of substrate oligonucleotide required the use of an excess of [^{32}P]-ATP (32 fmole/well) to ensure complete labelling. Substrate (12 fmole/well) was added to each microtitre well in the form of a labelling reaction mix that also contained an excess amount of [^{32}P]-ATP, it was not clear whether the inconsistencies detected were due to non-specific binding of the unincorporated excess [^{32}P]-ATP to the microtitre well or to inconsistent substrate binding.

To investigate this problem an equivalent concentration of [^{32}P]-ATP was added to eight separate wells either alone (0.02μl in a 100μl PBS/BSA) or in the form of radiolabelled substrate (+ unincorporated). After a two hour incubation period at room temperature, wells were washed five times in PBS and binding measured in a scintillation counter (1209 Rackbeta liquid scintillation counter, LKB Wallace; counting efficiency 59%). Bound radioactivity in the substrate wells represented 42% of labelled oligonucleotide added (5.1 fmole bound vs. 12 fmole added) and 15.9% of all [^{32}P]-ATP added. The coefficient of variance (CV) of counts measured in wells containing substrate was 16.5%. The residual counts measured in the [^{32}P]-ATP only wells represented 0.02% of all [^{32}P]-ATP added to each well and 0.05% of substrate bound (Table 5.1).

Table 5.1 Comparison of radioactively labelled oligonucleotide and [^{32}P]-ATP binding to microtitre plates.

Substrate Added	Mean cpm / well ± sd (range)	Proportion of [^{32}P]-ATP added / well (%)
Oligonucleotide	45 149 ± 7 458 (34 816 – 54 287)	15.9
[^{32}P]-ATP	45 ± 32 (21 - 114)	0.02

Since only 0.02% of initial [^{32}P]-ATP added became bound it was concluded that inter-well variation in binding was not due to non-specific binding of unincorporated [^{32}P]-ATP. The problem of inconsistent substrate binding was addressed by pre-blocking each well with 5% BSA-PBS for 2 hours at room temperature. The introduction of this step had the effect of doubling substrate binding and markedly improving substrate binding consistency when compared to non-blocked wells (CV 0.4% in blocked wells vs. 19.8% in non-blocked wells) (Table 5.2). The binding of [^{32}P]-ATP remained extremely low in blocked wells (30 ± 13 mean cpm / well, range 17 – 57).

Table 5.2 Substrate binding in blocked and non-blocked microtitre wells.

	Oligonucleotide Substrate Binding	
	Blocked Wells	Non-blocked Wells
Mean cpm / well ± sd (range)	17608 ± 63 (17536-17667)	8394 ±1660 (6725-9857)
C.V. (%)	0.4	19.8

A pre-blocking step was therefore introduced as standard into the preparation of microtitre plates. The optimum concentration of BSA-PBS for blocking was investigated by comparing the amount and variability of binding of substrate in wells blocked with 0%, 0.5%, 1%, 3% and 5% BSA-PBS. The CV of substrate binding was 14.7% in the non-

blocked plates and 4.1 to 6.8% in the blocked plates; twice as much substrate bound in blocked compared to non-blocked wells (Figure 5.2). There was no difference in the amount or variability of binding between the different concentrations of blocking agent therefore 0.5% BSA-PBS was chosen.
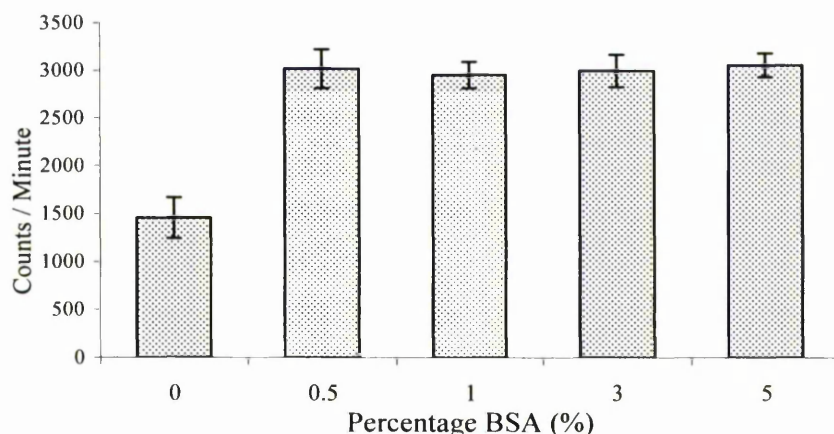
Figure 5.2 Mean oligonucleotide binding / well (cpm) in non-blocked and pre-blocked microtitre wells (using 0.5, 1.0, 3.0 and 5% BSA-PBS).

5.1.4 Screening of Peripheral Blood Mononuclear Cells

Initial analysis of MGMT in peripheral blood mononuclear cells (PBMC) showed ten fold less DNA repair activity than would have been expected according to data from previous studies in the same laboratory (Margison, Heighway et al. 2005). The reason for this discrepancy was investigated by evaluating the method of PBMC sample preparation (see Section 2.2.9.1 and Section 2.4.6.3). This process involved the freezing of cells in a PBS suspension followed by rapid defrosting when samples were due to be analysed. After centrifugation, the supernatant was discarded and the remaining pellet sonicated and assayed. However, assays of the discarded supernatant demonstrated that its MGMT activity was significantly greater than that of the pellet.

To investigate this further 12 PBMC samples were defrosted, vortex mixed and split into two equal aliquots. One aliquot ('nuclear extract') was processed using the previously described method (Section 2.4.6.3). The second aliquot ('whole cell extract') was sonicated after the addition of 0.1 volumes of 10x buffer I prior to following the method

previously described (Section 2.4.6.3). DNA and protein concentrations were calculated for the pellet and supernatants of all 24 samples (Table 5.3). The nuclear extract contained on average 98.6% (range 94.6 to 99.9%, CV = 1.5%) of the whole sample DNA and 12.4% (range 7.6 to 17.4%, CV = 26.0%) of the whole sample protein concentration.

Table 5.3 Comparison of PBMC nuclear and supernatant / whole-cell extracts: MGMT activity, DNA and protein concentration.

Measure		PBMC Samples (n = 12)	
		Nuclear Extract	Supernatant or Whole for MGMT
DNA	Mean Concentration (µg/ml) ± sd	38.3 ± 14.7	0.6 ± 0.6
	% of whole sample (range)	98.6 (94.6 – 99.9)	1.4 (0.1-5.4)
Protein	Mean Concentration (mg/ml) ± sd	0.65 ± 0.26	4.50 ± 1.04
	% of whole sample (range)	12.4 (7.6-17.4)	87.6 (82.6-92.4)
MGMT	Mean Activity (fmole/ml) ± sd	39.5 ± 23.9	712.1 ± 229.5
	% of total sample (range)	5.5 (2.5 – 14.1)	100

Despite efforts to ensure homogeneity and equally separate the thawed PBMC aliquots, the mean DNA concentration of samples after sonication was consistently greater in the nuclear extract compared to the whole cell extract (38.3 ± 14.7 vs. 24.9 ± 10.6µg/ml). Samples were diluted in buffer I to a DNA concentration of 10µg/ml; nuclear extracts were on average diluted 1.5 times that of the whole cell sample. MGMT activity in this instance was therefore calculated for the whole sample (fmole/ml) and not relative to DNA concentration (fmole/µgDNA). Using this method, MGMT activity in the nuclear extract represented only 5.5% (range 2.5 – 14.1%) of MGMT activity measured in the whole-cell extract (Table 5.3). Measured whole sample MGMT activity was therefore similar to previously reported values for PBMCs (Margison, Heighway et al. 2005).

5.1.5 Sonication Conditions

5.1.5.1 Sonication Time

Sonication had been the method used by the Carcinogenesis Group for the preparation of cellular samples for MGMT assays for many years. However, it was not known if sonication had an effect on the cleavage activity of MPG. Sonication introduces a time dependent increase in temperature of the sample being processed. To illustrate this 1ml aliquots of ice cold buffer IBSA were sonicated for increasing amounts of time and the temperature change recorded, using a thermocouple thermometer (Figure 5.3). Sample temperature increased by $1.5^{\circ}C$ for each additional second of sonication. The effects of increasing sonication time on the amounts of DNA and protein extracted from MCF-7 cell pellets and the associated effects on MPG and MGMT oligonucleotide cleavage activity were investigated (Figure 5.4).

The amount of protein in the extract did not increase with time, suggesting that sonication for 5s was effective in releasing almost all protein. However, the amount of DNA extracted increased by 25% from 5s to 10s. There was a further increase in DNA extraction between 10s and 20s but this was limited to only 5%. To investigate the effect of sonication time on cleavage activity a pellet of MCF-7 cells was split into four equal aliquots and each was sonicated for different times (5, 10, 15 and 20s). Increasing sonication time resulted in a linear reduction of MPG activity (0.5 fmole/µgDNA/hr for each second of sonication). MGMT activity was relatively stable from 5s to 15s (5% reduction in activity), before a 15% drop in activity was seen when sonication time was increased to 20s (Figure 5.4). A sonication time of 10s was chosen as standard for this assay. This optimised DNA extraction without significantly effecting MGMT activity, the loss of MPG activity of 10% from 5s to 10s was considered acceptable.
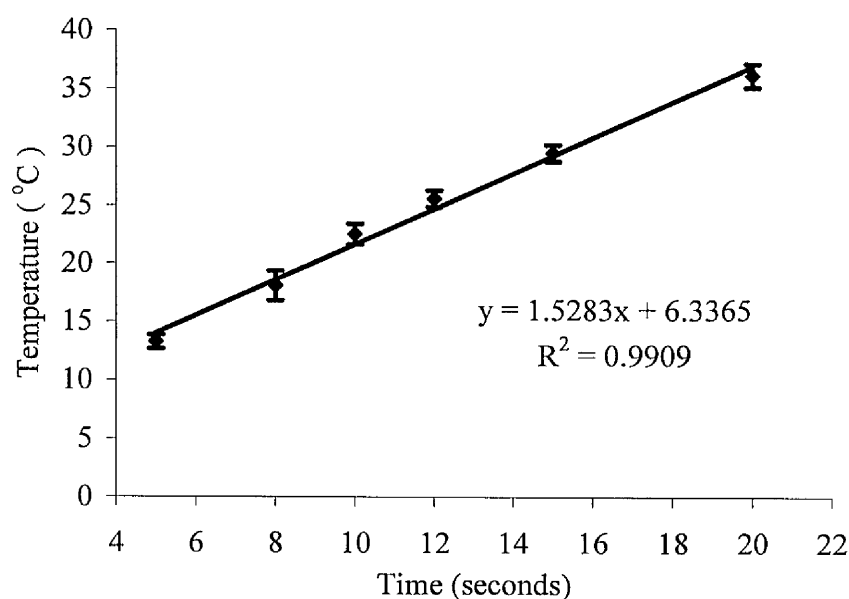
Figure 5.3 Sample (IBSA) temperature with increasing sonication time.
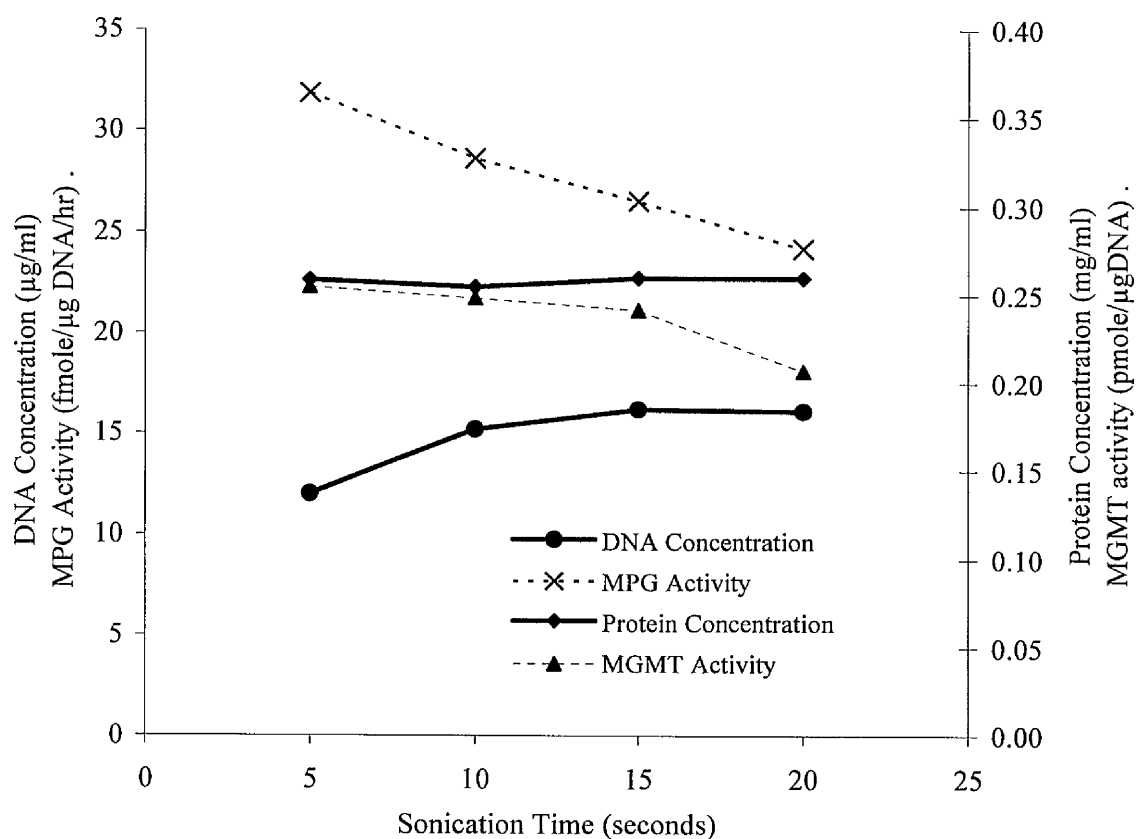
Figure 5.4 The effect of increasing sonication time of MCF-7 cells on protein and DNA extraction and the activity of MGMT and MPG.

5.1.5.2 Sonication Buffer

Assays using the standard laboratory method produced measurable oligonucleotide cleavage activity for both MGMT and MPG in MCF-7 cell pellets (Table 5.4). However, MPG cleavage represented less than double background (Figure 5.5B). Attempts to improve the assay were therefore made.

The first modification of the standard laboratory method was to sonicate cell pellets in 1ml buffer K (common buffer taken from study by Kreklau *et al*: 25mM HEPES-KOH, 150mM KCl, 0.5mM DTT, 0.5mM EDTA, 1% glycerol, pH 7.8 + leupeptin 10mg/ml) (Kreklau, Limp-Foster et al. 2001), instead of 1ml buffer I (50mM Tris-HCl, pH8.3, 1mM EDTA, 3mM DTT + leupeptin 10mg/ml). Buffer I sonicates were diluted 1 in 10 into cleavage buffer; buffer K sonicates were diluted 1 in 10 in fresh buffer K or in buffer KBSA, before transfer to the 96-well microtitre plate (100µl/well). All assays were performed in triplicate (Figure 5.5). MCF-7 cellular extract sonicated in Buffer K or buffer KBSA showed a doubling in MGMT repair activity when compared to the standard method. However, MPG activity was 50% lower in extract sonicated in buffer K and slightly lower in buffer KBSA compared to standard assay conditions. The amount of DNA extracted from the sonication of MCF-7 cell pellets was double in buffer I compared to buffer K (Table 5.4). All extracts were diluted to 10µg/ml prior to the cleavage assay; the buffer I samples were therefore diluted twice as much as the buffer K samples. The raised MGMT activity level, when expressed in terms of DNA, was possibly due to reduced efficiency of DNA extraction when samples were sonicated in buffer K. It was therefore decided not to use buffer K, but to continue with buffer I and maximise sample extraction.

Table 5.4 Comparison of Buffer I, K and K+BSA using MCF-7 cell extracts and oligonucleotide cleavage assays for MGMT and MPG.

Buffer Used	Supernatant DNA Concentration (µg/ml)	Oligonucleotide Cleavage Activity ± sd (CV %)	
		MGMT (fmole/µgDNA)	MPG (fmole/µgDNA/hr)
I	61.4	16.7 ± 0.6 (3.3)	1.9 ± 0.1 (7.1)
K	28.4	37.2 ± 1.5 (4.0)	0.9 ± 0.1 (8.7)
K+BSA		37.9 ± 4.2 (11.0)	1.7 ± 0.1 (5.4)

(a)   MGMT Assay
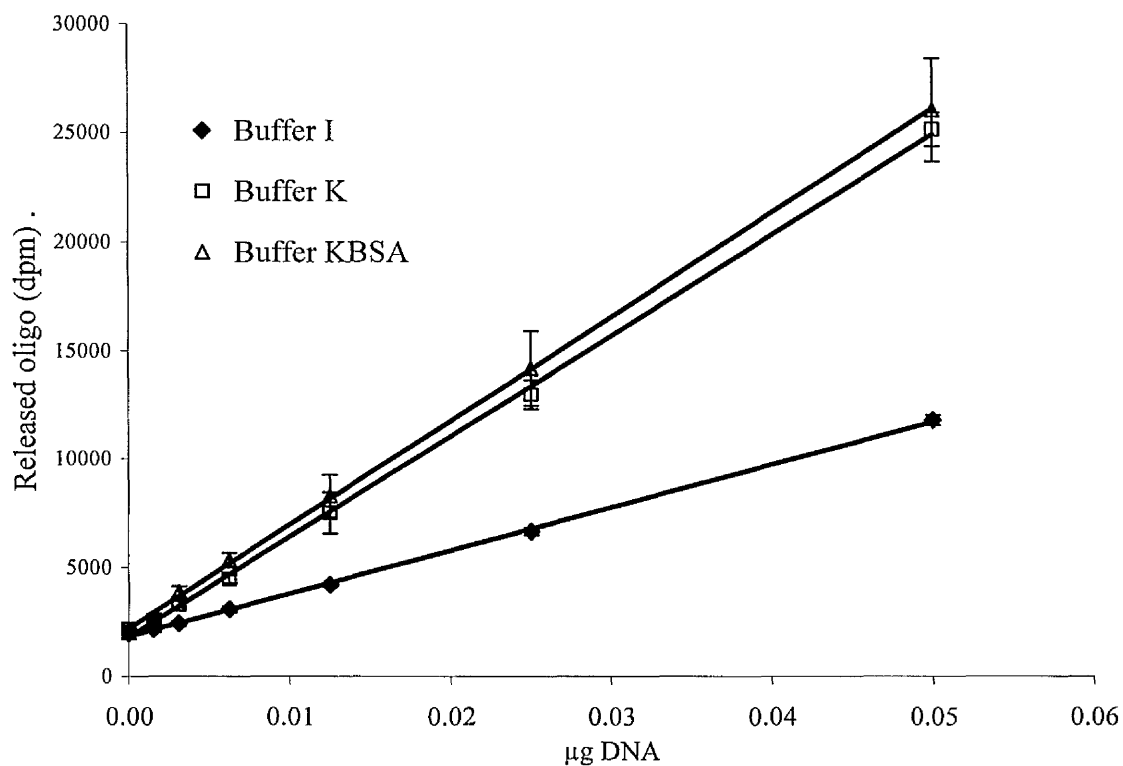
(b)   MPG Assay
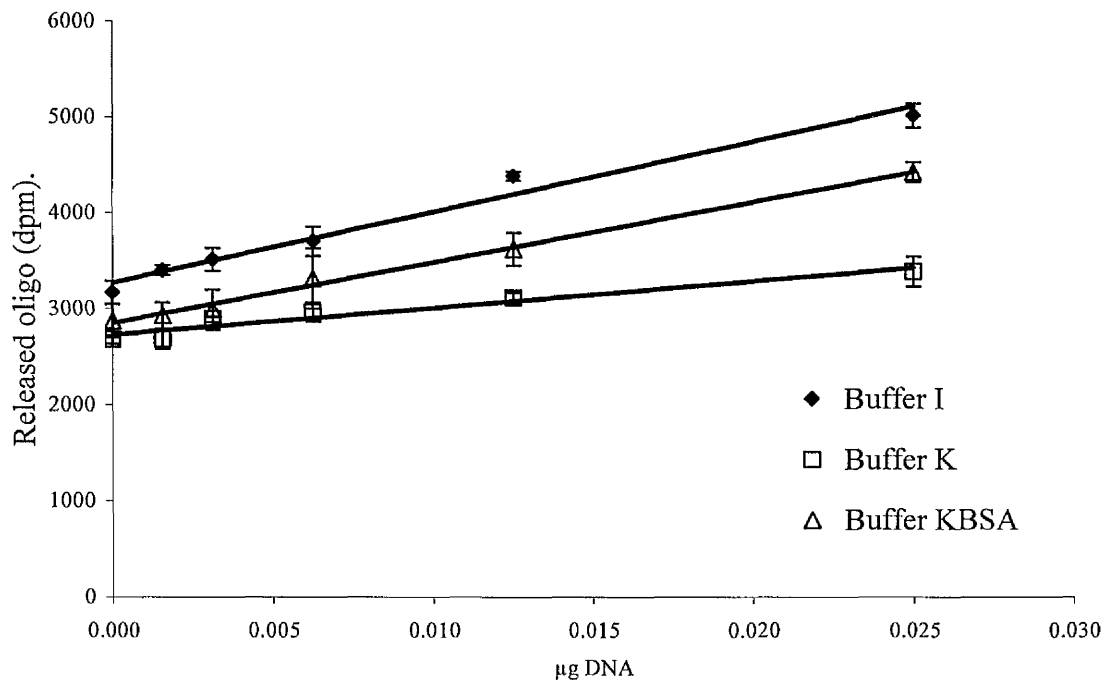
Figure 5.5a+b Oligonucleotide cleavage activities for a) MGMT and b) MPG using sonication buffer I, K or K+BSA (MCF-7 cell extract).

175

5.1.6 MPG Cleavage Buffer

The cleavage buffer used for the MPG assay initially consisted of 25mM Tris-HCl (pH 7.6), 50mM KCl and 5mM EDTA. MPG is a monofunctional glycosylase, the removal of the εA adduct would have created an AP site which *in vivo* would then be cleaved by APE-1, the initiator of the second stage of BER. However, APE-1 is Mg^{2+} dependent and would therefore have been inactive in the standard assay buffer due to the Mg^{2+} chelating properties of EDTA (Wilson 2005). Without AP site cleavage in the assay it was postulated that repair of the εA adduct would not necessarily have been followed by oligonucleotide cleavage and release of the radiolabelled product. Removal of the εA adduct could therefore have been underestimated in the original buffer.

To address this problem it was decided to activate APE-1 in the assay either by removing EDTA or by adding Mg^{2+} to the cleavage buffer. It was postulated that not only would APE-1 be activated under these circumstances but there could also be activation of exo and endonucleases causing non-specific oligonucleotide cleavage (Yang, Lee et al. 2006). A control double stranded oligonucleotide was introduced having the same sequence as the MPG substrate oligonucleotide but with a normal base instead of the adduct. This was assayed alongside the usual substrate oligonucleotide to investigate the extent of non-specific cleavage. The removal of EDTA from the cleavage buffer resulted in a large amount of non-specific cleavage of both substrate and control oligonucleotides, this was therefore abandoned.

The effect of Mg^{2+} in the cleavage buffer was assessed using an MCF-7 cell extract. MPG and control cleavage activity was determined at incrementally increasing Mg^{2+} concentrations (0 to 10mM) (Figure 5.6). There was an increase in MPG cleavage activity in the MCF-7 cell extract. An increase in control oligonucleotide cleavage was also seen suggesting the activation of non-specific nucleases. The effects were greatest at concentrations of Mg^{2+} above 5mM and reached a plateau at 10mM.
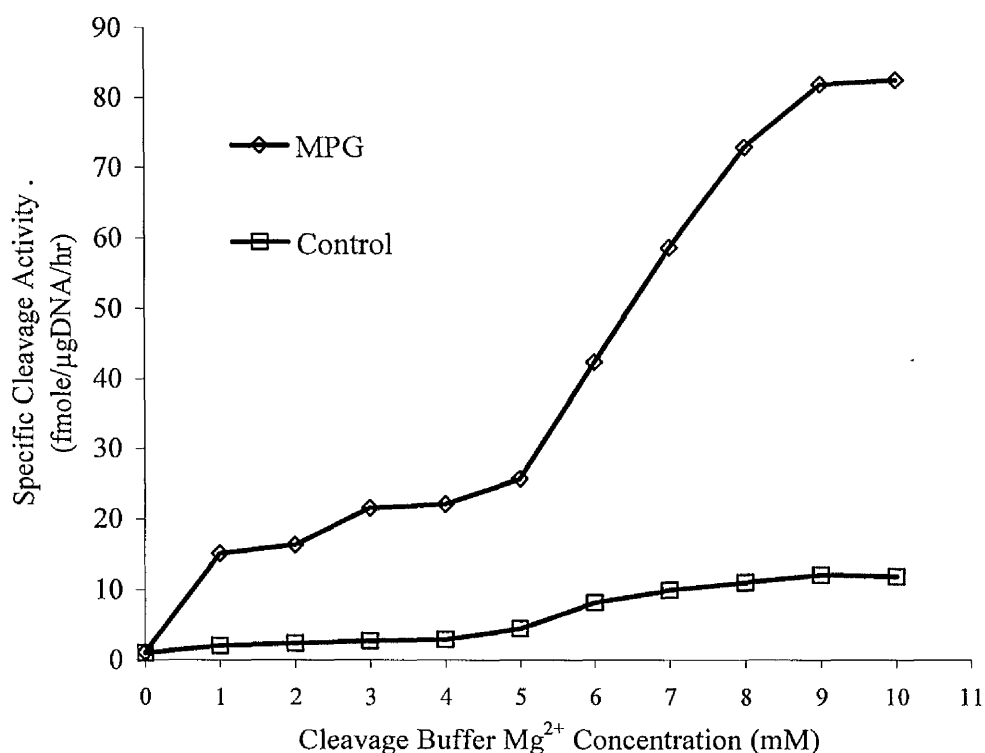
Figure 5.6 Cleavage buffer Mg^{2+} concentration and MPG and control cleavage activity (MCF-7 cell extract).

The effect of adding increasing amounts of pure APE-1 protein on the cleavage activity of MPG was investigated in an MCF-7 cell extract using Mg^{2+} cleavage buffer (10mM). Increasing additional APE-1 protein had no impact on MPG cleavage activity suggesting APE-1 was not a rate limiting factor in the cleavage reaction (Figure 5.7).
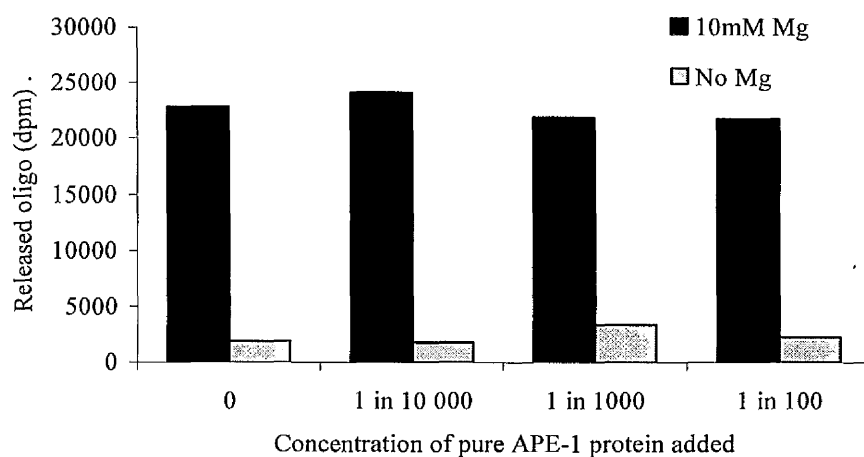
Figure 5.7 The addition of APE-1 protein and MPG cleavage activity in an MCF-7 cell extract.

177

5.1.7 Kinetics

The effect of increasing incubation time on the activity of both MGMT and MPG was determined using MCF-7 cell extract. Cleavage of both the MPG and control oligonucleotide was linear over a 2.5hr period (Figure 5.8). A time of two hours was chosen for subsequent assays. The MGMT assay was required to go to completion for an accurate measure of activity. A kinetics experiment showed that the reaction went to greater than 98% completion after a time of four hours. For convenience four hours was therefore chosen as the length of time for the assay (Figure 5.9). Assays performed over 6 or more hours showed a marked reduction in activity. The assays were performed at 37°C and it was thought that radioactivity may have been lost through drying to the side of wells as the reaction buffer evaporated.
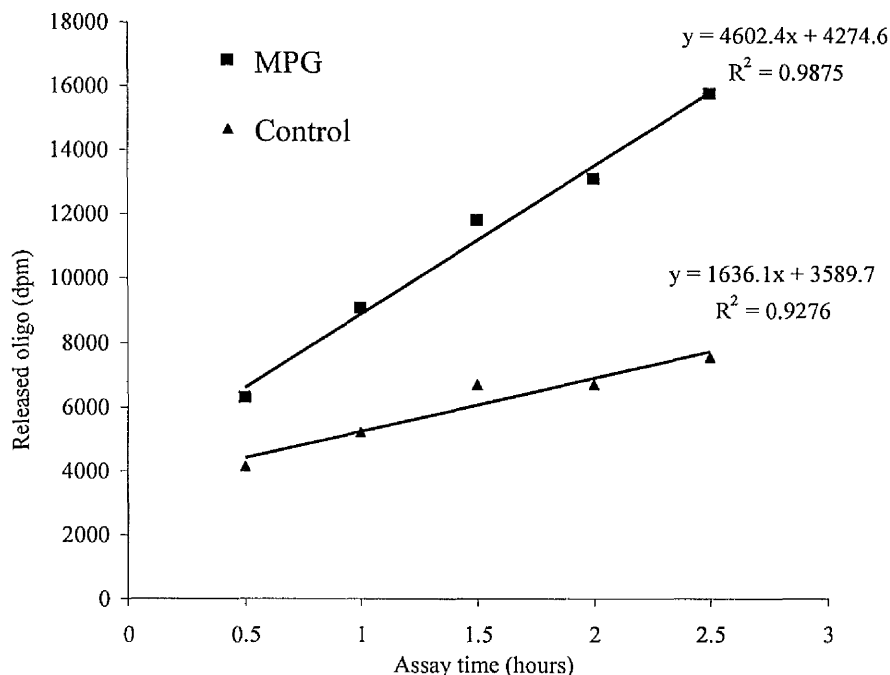
Figure 5.8 Kinetics of the MPG assay for substrate and control oligonucleotide (MCF-7 cell extract).
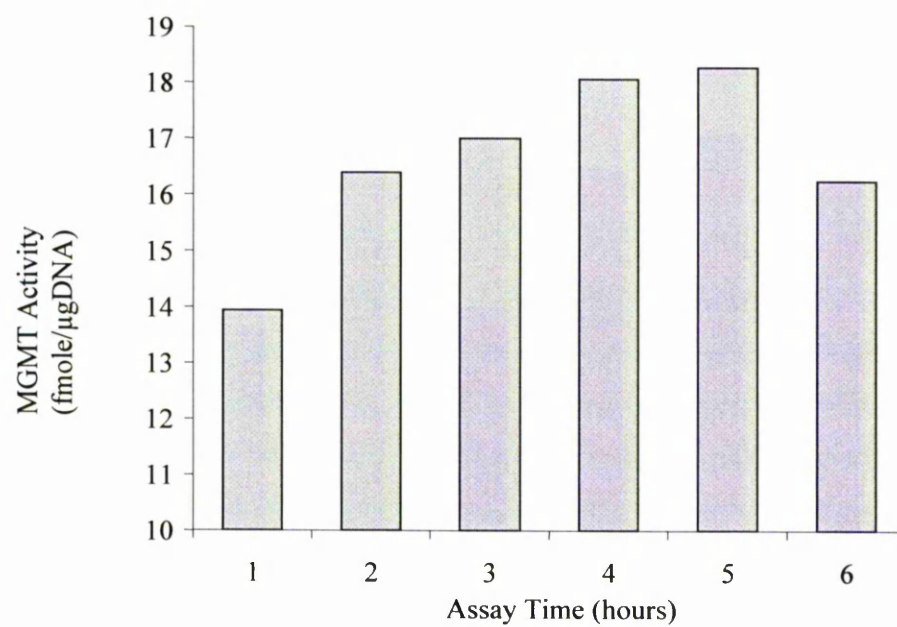
Figure 5.9 Kinetics of the MGMT assay using MCF-7 cell extract.

5.2 MGMT and MPG Assay Validation

5.2.1 Determination of Precision and Lower Limit of Quantitation

The lower limit of quantitation (LLoQ) was defined as the lowest amount of cleavage measurable with necessary precision above the limit of detection (3 times the standard deviation of the negative control value). Both the MPG and MGMT assays required linearity ($R^2 \geq 0.985$) of at least 3 points to produce a gradient from which activity was calculated. The LLoQ for specific activity of the MGMT assay was therefore four times higher than the LLoQ for the lowest measurable amount of cleavage because samples were double diluted. The LLoQ for the MPG assay was therefore 2.2 times the LLoQ for the lowest measurable amount of cleavage because samples were diluted 1.3 fold. To determine the limits of detection and quantitation for both the MPG and MGMT assays, repair activities were measured in reducing volumes of a sonicated MCF-7 sample in sextuplicate (n = 6). The limit of detection for the MGMT assay was 0.03 fmole. The lower limit of quantitation for MGMT specific activity was 0.22 fmole/µg DNA. The limit of detection for the MPG assay was 0.09 fmole/hr and the lower limit of quantitation for MPG specific activity was 0.71 fmole/µgDNA/hr (Figure 5.10a+b).
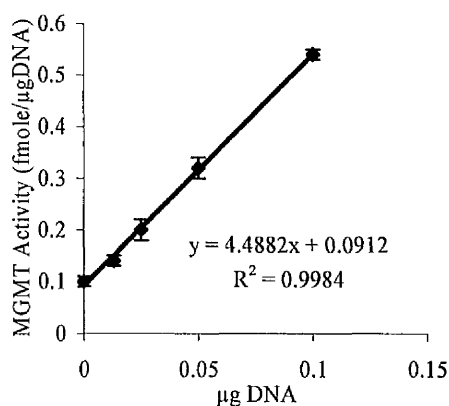
(a) MGMT Activity　　　　　**(b) MPG Activity**
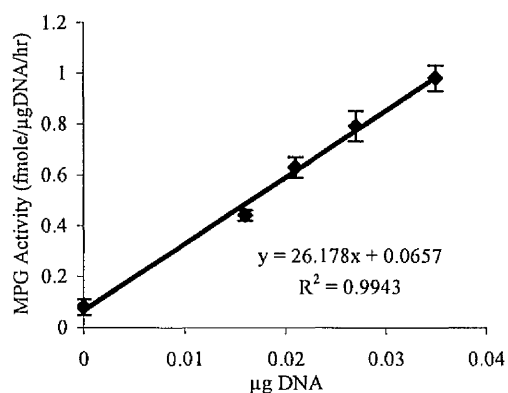
Figure 5.10 LLoQ for (a) MGMT and (b) MPG cleavage activity in an MCF-7 extract. Each point is the mean ± sd of sextuplicate determinations.

5.2.2 Inter-Assay Variability and the Determination of Acceptance Limits for the Quality Control Sample

To investigate inter-assay variability both MPG and MGMT activity was measured in MCF-7 cellular extracts on eight separate occasions. Measurements of activity were performed on eight different MCF-7 cell pellets, carried out on separate days, using substrate labelled on different days. The mean MPG activity was 22.1 ± 5.6 fmole/μgDNA/hr. The mean MGMT activity was 16.8 ± 3.7 fmole/μgDNA. The measurement of study samples would include an MCF-7 cellular extract acting as a quality control. The result of an assay would be defined as acceptable if the MCF-7 extract value fell between 13.1 – 20.5 fmole/μgDNA for MGMT (assay CV 22%) and 16.5 – 27.7 fmole/μgDNA/hr for MPG (assay CV 25%). Therefore the precision of measurement of MGMT and MPG was 22% and 25% respectively. This is acceptable since precision of biological assays are normally set at 25% (Thorpe, Wadhwa et al. 1997).

5.2.3 Assay Specificity

5.2.3.1 Specificity of MGMT Cleavage

To investigate the specificity of the MGMT repair assay MCF-7 cellular extract and IBSA only as negative controls were incubated both with (+) and without (-) the addition of the *Pst*1 restriction enzyme. In *Pst*1 (+) wells, there was a linear increase in MGMT cleavage activity with increasing extract DNA concentration. No cleavage above background was seen when extract was incubated in *Pst*1 (-) wells. However there was a significant difference in values measured from *Pst*1 (-) and (+) negative control wells. The increase in counts from *Pst*1 (+) wells in the absence of extract suggested the presence of non-specific cleavage (Figure 5.11a).

The same experiment was repeated using an MGMT control oligonucleotide (Figure 5.11b). This had the same sequence as the MGMT substrate oligonucleotide but a normal base replaced the adduct (O^6-meG replaced by guanine). In *Pst*1 (-) wells no discernable cleavage was detected and in *Pst*1 (+) wells a mean of 99.3% of control oligonucleotide was cleaved (80 408 ± 3460 dpm / well cleaved vs. 80 980 ± 1417 dpm / well control substrate bound). The degree of control oligonucleotide cleavage was independent of extract concentration and was also the same in negative control wells. The assay

conditions used for *Pst*1 were therefore appropriate as almost complete well substrate cleavage was achieved. This experiment also provided indirect evidence that the presence of O^6-meG prevented cleavage by *Pst*1. There was approximately an 80 fold difference in released counts between *Pst*1 (+) and (-) negative control wells between substrate and control oligonucleotides.

(a)   MGMT Substrate Oligonucleotide
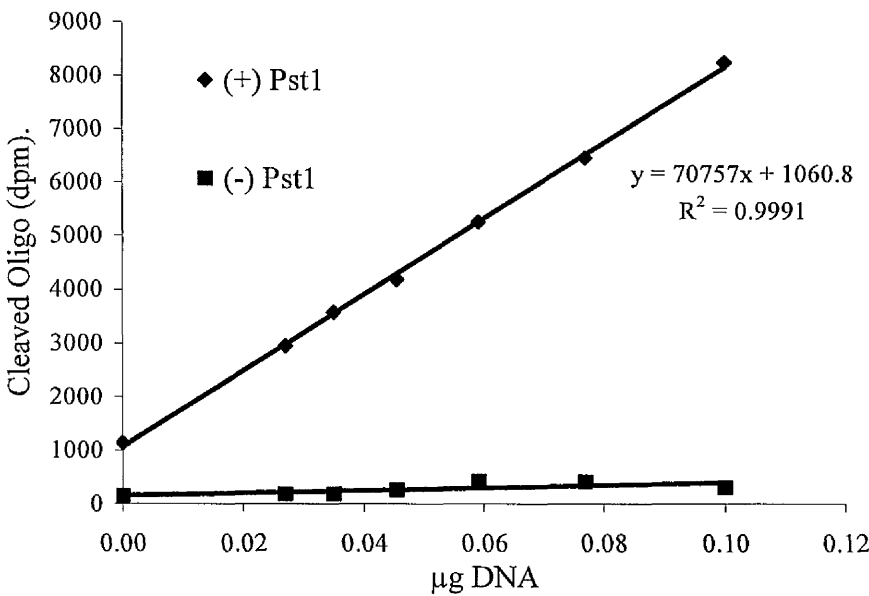
(b)   MGMT Control Oligonucleotide
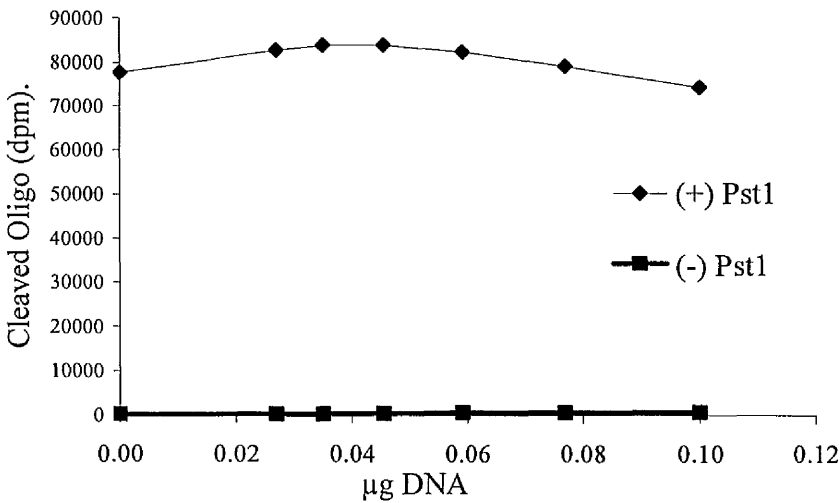
Figure 5.11a+b Comparison of cleavage activity with (+) and without (-) the restriction enzyme *Pst*1 on (a) MGMT substrate and (b) MGMT control oligonucleotides (MCF-7 cellular extract).

5.2.3.2 Specificity of MPG Cleavage

The addition of Mg^{2+} had resulted in increased cleavage of MPG substrate. The activity of pure MPG and APE-1 both individually and combined were investigated to establish if the increased cleavage was due to either enzyme alone or in combination. Pure MPG protein (1 in 200 000) was assayed with and without APE-1 protein (1 in 100) in the Mg^{2+} cleavage buffer (10mM). Cleavage of MPG substrate was increased by a factor of 21 fold in the presence of APE-1 when compared to the MPG protein alone (Figure 5.12). There was only a marginal increase in dpm released by activated APE-1 compared to no APE-1 when MPG was absent (Figure 5.12), suggesting that APE-1 alone was not responsible for significant oligonucleotide cleavage.
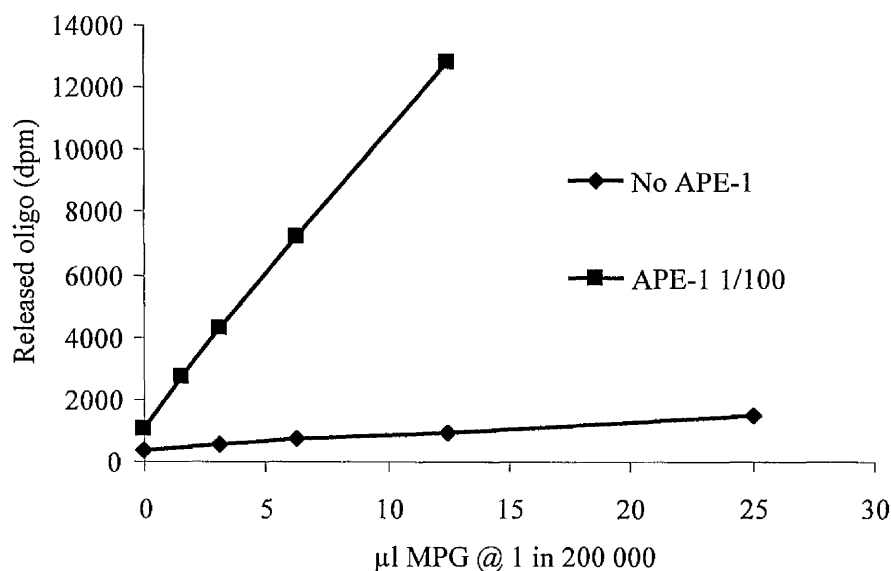
Figure 5.12 MPG substrate cleavage in magnesium cleavage buffer ± APE-1 (1/100).

The addition of Mg^{2+} to the cleavage buffer had increased MPG substrate cleavage but also increased non-specific cleavage in MCF-7 cellular extracts (Figure 5.6). To investigate the extent and nature of non-specific cleavage in PBMCs, ten whole and ten nuclear extracts were assayed using MPG and control oligonucleotide substrates. The mean whole PBMC MPG repair activity was 19.4 ± 5.5 (range 14.3 – 30.5) fmole/µgDNA/hr and 10.9 ± 3.6 (range 7.1 – 18.6) fmole/µgDNA/hr for the MPG control. Control substrate cleavage represented on average 56% of MPG substrate cleavage (range 43.7 – 71.0%) (Figure 5.13a).

The mean nuclear PBMC MPG repair activity was 3.88 ± 1.18 (range 1.92 – 5.60) fmole/µgDNA/hr and 0.34 ± 0.04 (range 0.28 – 0.40) fmole/µgDNA/hr for the MPG control. Control substrate cleavage represented on average 9.5% of MPG substrate cleavage (range 5.5 – 16.3%) (Figure 5.13b). The CV of control oligonucleotide cleavage was 33.4% for whole PBMC extracts and 11.9% for the nuclear PBMC extracts.

To determine the size and nature of products cleaved by PBMC extracts in magnesium cleavage buffer when assayed on MPG and control oligonucleotide substrates, a 25µl aliquot of post-assay product was mixed with 5µl 6x non-denaturing loading buffer and separated using PAGE before visualisation on a STORM phosphorimager. It was initially determined that pure MPG / APE-1 cleaved a 5mer product on MPG substrate with no cleavage seen when using the control (Figure 5.14). This finding was consistent with specific cleavage of the εA adduct. The MPG / APE-1 cleavage product was therefore subsequently used as a size marker. Analysis of cleavage products from assays of PBMC nuclear and whole cellular extracts confirmed that both extracts produced the same major cleavage product as pure MPG / APE-1 (Figure 5.15). There was evidence of additional products seen in both assays from both substrate and control oligonucleotides; these included double and single stranded oligonucleotides, no other discernable bands were seen. The degree of single and double stranded oligonucleotide released was judged to be greater from control oligonucleotide wells. This may have accounted for some of the non-specific release from control oligonucleotide wells seen in all assays. An additional product was also detected in control assays of whole PBMC extract that was the same size as the pure MPG / APE-1 product; this was not detected when nuclear PBMC extract was assayed on control oligonucleotide (Figure 5.15).

(a) Whole PBMC Extracts
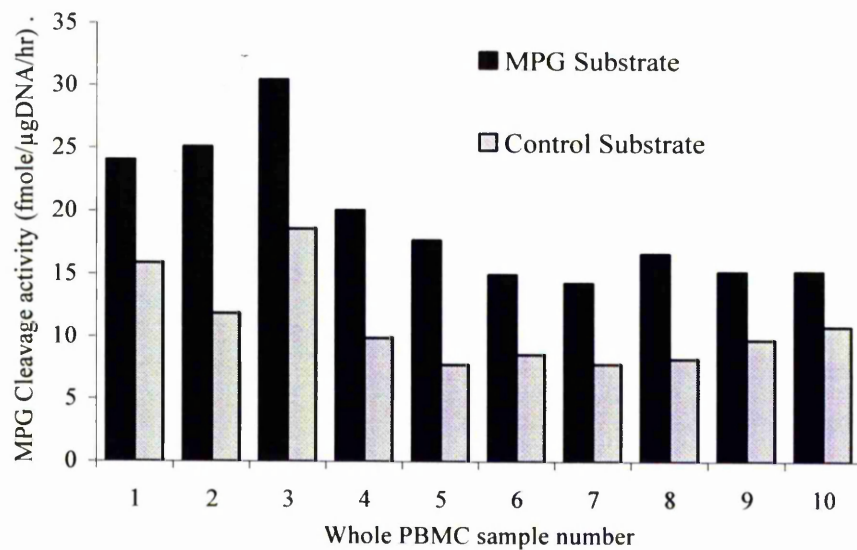
(b) Nuclear PBMC Extracts
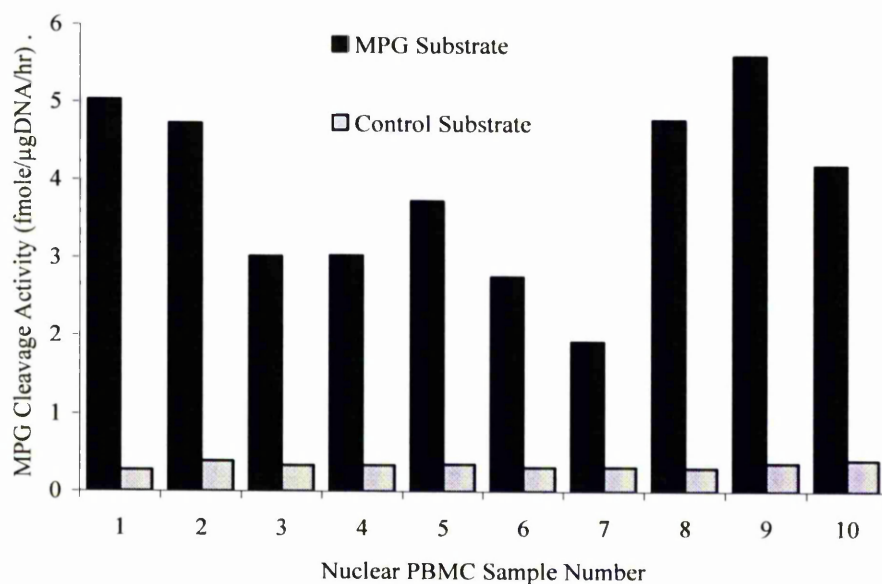
Figure 5.13a+b Comparison of MPG substrate and MPG control oligonucleotide cleavage using PBMC (a) whole and (b) nuclear extracts.
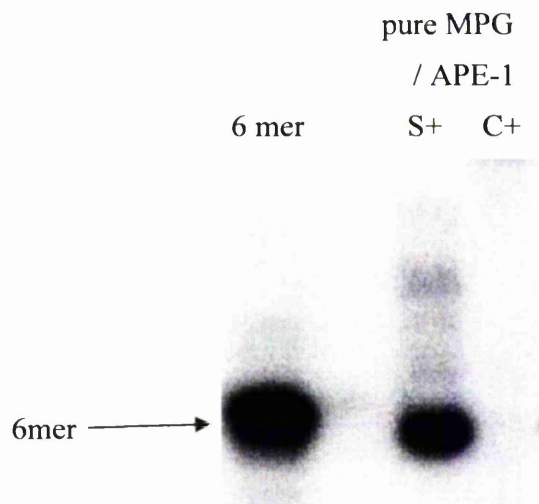
Figure 5.14 Visualisation and sizing of the product of pure MPG / APE-1 cleavage of MPG substrate and control oligonucleotide.

S = substrate oligonucleotide, C = control oligonucleotide, (+) = pure enzyme added
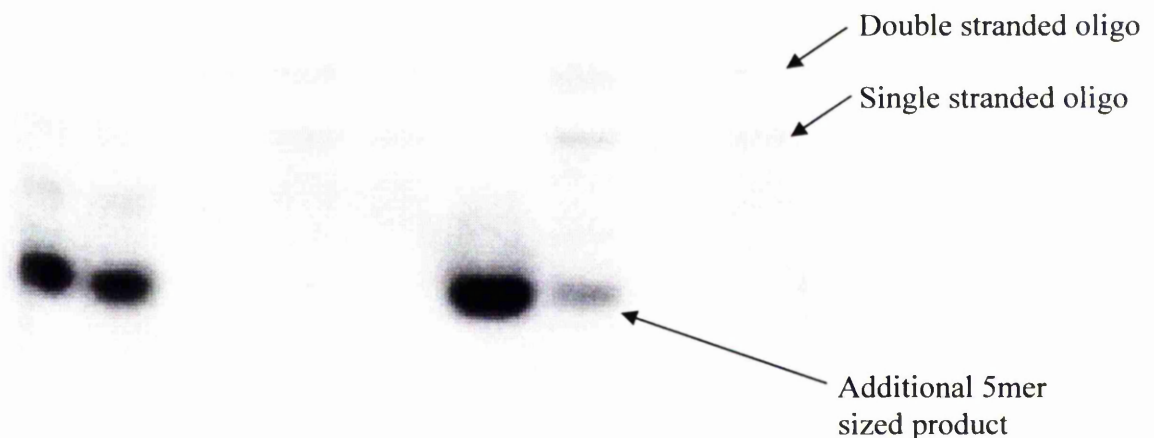
Figure 5.15 Separation (PAGE) and visualisation (STORM phosphorimager) of MPG assay products using: pure MPG/APE-1 and nuclear / whole PBMC samples.

S = substrate oligonucleotide, C = control oligonucleotide,

(+) = extract added, (-) = IBSA added with no extract.

The cleavage of control oligonucleotide by whole PBMC extract was further investigated. Firstly, the observation was repeated after incubation of extract from three different PBMCs on control oligonucleotide (Figure 5.17a). Secondly, the amount of 5mer produced increased in proportion to the amount of extract incubated (Figure 5.17b). A second control oligonucleotide was ordered and assayed to rule out contamination by substrate oligonucleotide. The same finding was seen after incubation of whole PBMC extract. It was not apparent why a 5mer sized product should be cleaved from the MPG control oligonucleotide, the assay was repeated using the MGMT control oligonucleotide and no equivalent band was seen (Figure 5.16). This suggested that the cleavage of the MPG control may have been sequence related, but this idea was not supported by the lack of cleavage seen with pure MPG / APE-1, or related to other nucleases present in the whole cell extract.

Pure MPG   Whole
/ APE-1   PBMC Extract
S1+   C2+ C2-
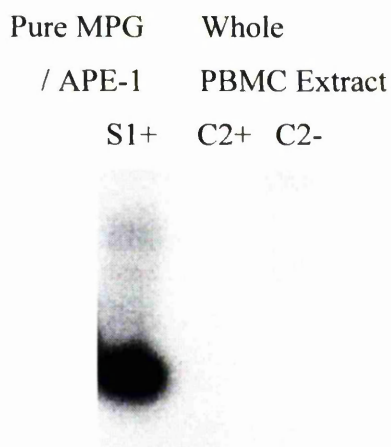
Figure 5.16 Products of whole PBMC extract assayed on MGMT control oligonucleotide.

S1 = MPG Substrate Oligonucleotide, C2 = MGMT Control Oligonucleotide

(+) = extract added, (-) = IBSA added with no extract.

Analysis of the products from a further three nuclear PBMC extracts confirmed specific cleavage of the εA adduct by the presence of a single band at the same level as the pure MPG / APE-1 product (5mer) with no evidence of non-specific control oligonucleotide cleavage (Figure 5.18). Increasing concentrations of nuclear PBMC extract produced a linear increase in the 5mer product; there was no evidence of cleavage in the absence of sample extract (Figure 5.19). Measured increased cleavage from PBMC nuclear assays was therefore judged to be due to specific cleavage of the εA adduct.

(A) Three whole PBMC extracts

PBMC Whole Samples

(B) One whole PBMC extract
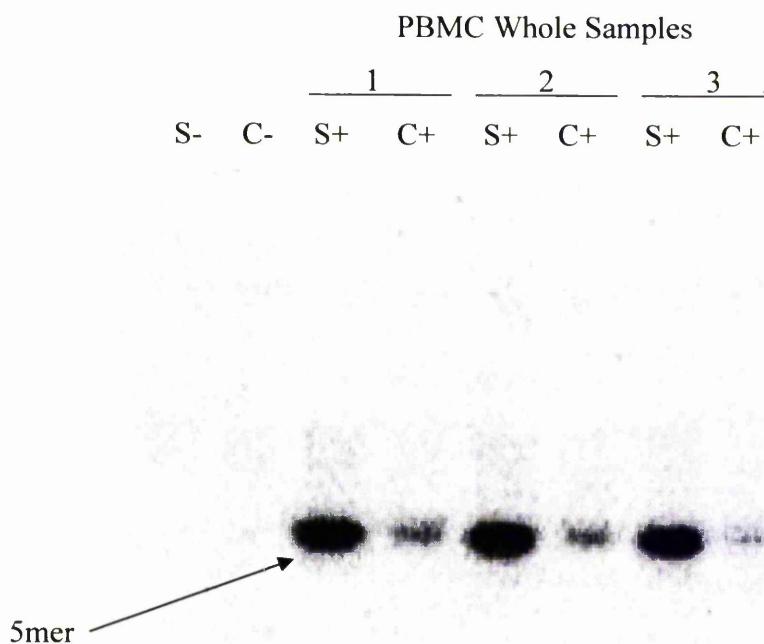
Relative amount of whole PBMC extract
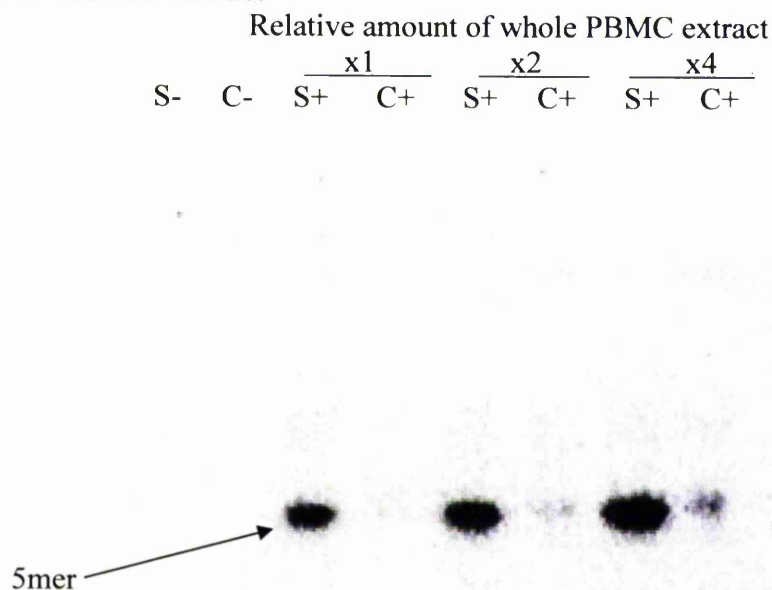
Figure 5.17a+b Products of whole PBMC cleavage assays on MPG and control substrate using: (A) three different samples and (B) increasing extract from one sample

S = substrate oligonucleotide, C = control oligonucleotide,

(+) = extract added, (-) = IBSA added with no extract.

Pure MPG / APE-1   Nuclear PBMC Samples

Figure 5.18 Products of MPG assay in three PBMC nuclear samples.

S = substrate oligonucleotide, C = control oligonucleotide, (+) = extract added.

Nuclear PBMC Relative Sample Concentrations

None   1   1.3   1.7   2.2   2.9
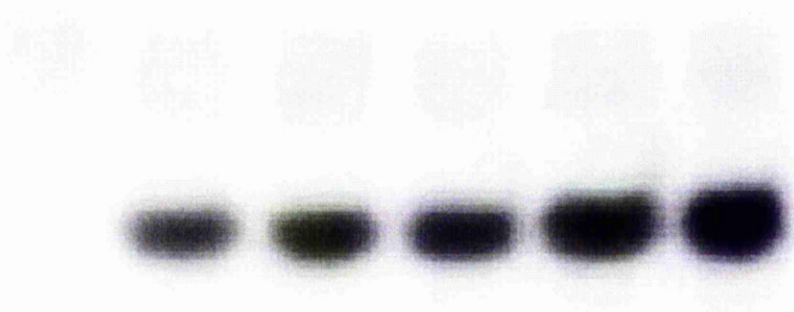
Figure 5.19 Intensity of 5mer with increasing nuclear PBMC extract concentration.

5.2.4 Correction of MPG Activity

The method of correcting MPG activity for the presence of non-specific cleavage in assays of PBMC extracts was investigated. A corrected MPG activity was calculated for each sample by subtracting the cleavage activity measured using the control substrate from the MPG substrate (Figure 5.20). MPG and corrected MPG values were significantly correlated ($R^2 = 0.63$; p = 0.007).

The chart shows a scatter plot of Corrected MPG Activity (fmole/μgDNA/hr) on the y-axis (0 to 14) against MPG Activity (fmole/μgDNA/hr) on the x-axis (0 to 35), with a linear trend line. The equation shown is $y = 0.4162x + 0.3978$ and $R^2 = 0.623$.
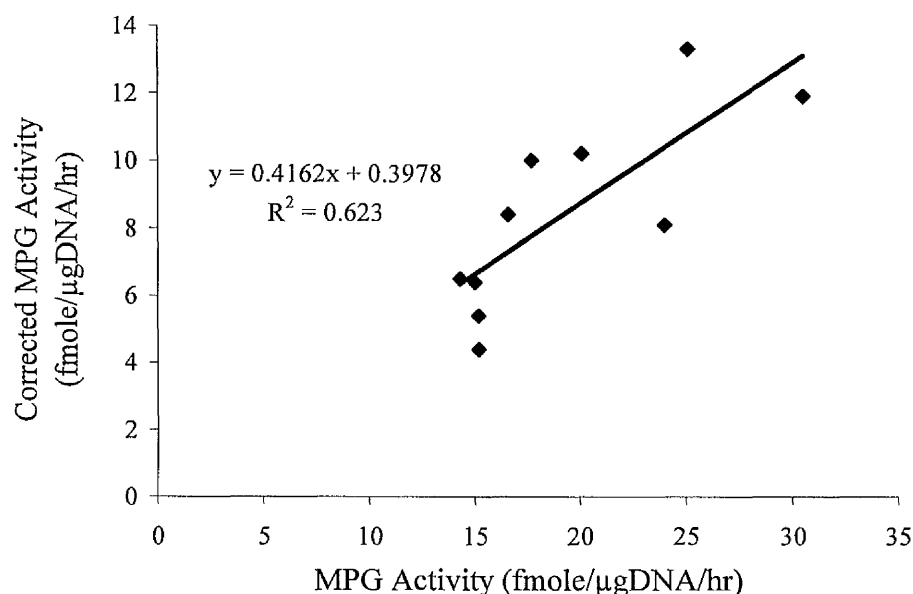
Figure 5.20 MPG plotted against corrected MPG activity in whole PBMC cellular extracts.

Correction of MPG activity by subtracting control oligonucleotide cleavage produced a mean corrected MPG activity of 8.5 ± 2.9 fmole/μgDNA/hr. The relationship between MGMT activity and MPG or corrected MPG repair activity in the 10 whole PBMC extracts was examined. The mean MGMT activity was 20.3 ± 7.7 fmole/μgDNA; MGMT activity was significantly correlated with both MPG (p = 0.001) and corrected MPG activity (p = 0.04) in whole PBMC samples (Figure 5.21a+b).

Although MPG and corrected MPG values were significantly correlated it was not known how appropriate or accurate this correction was; as the cleavage product was the same size in both control and MPG substrates (Figure 5.15) it was not possible to determine if the process resulting in control oligonucleotide cleavage was occurring in the MPG substrate.

(a) MPG vs. MGMT

$$y = 1.2165x - 3.3258$$
$$R^2 = 0.7446$$

MGMT Activity (fmole/µgDNA)

MPG Activity (fmole/µgDNA/hr)
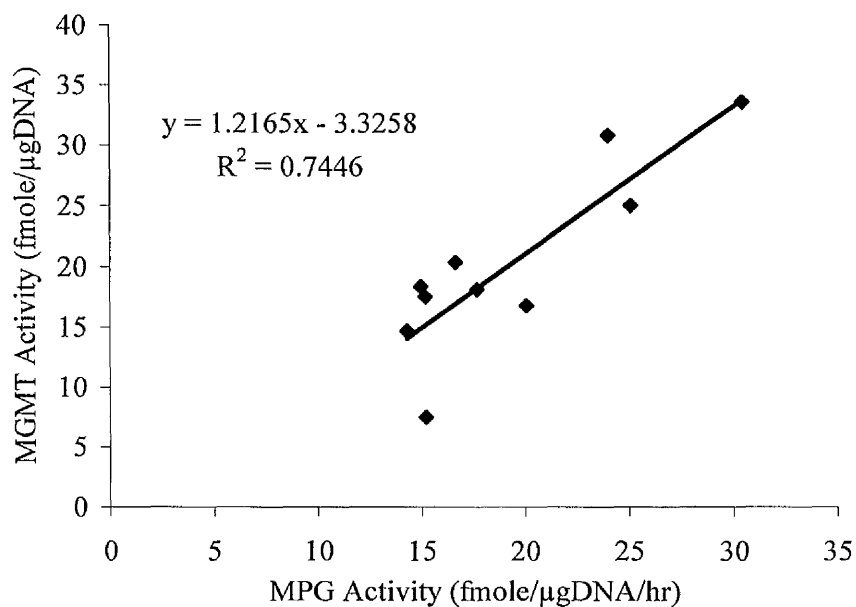
(b) Corrected MPG (MPG − control) vs. MGMT

$$y = 1.7742x + 5.2246$$
$$R^2 = 0.4364$$

MGMT Activity (fmole/µgDN/

Corrected MPG Activity (fmole/µgDNA/hr)
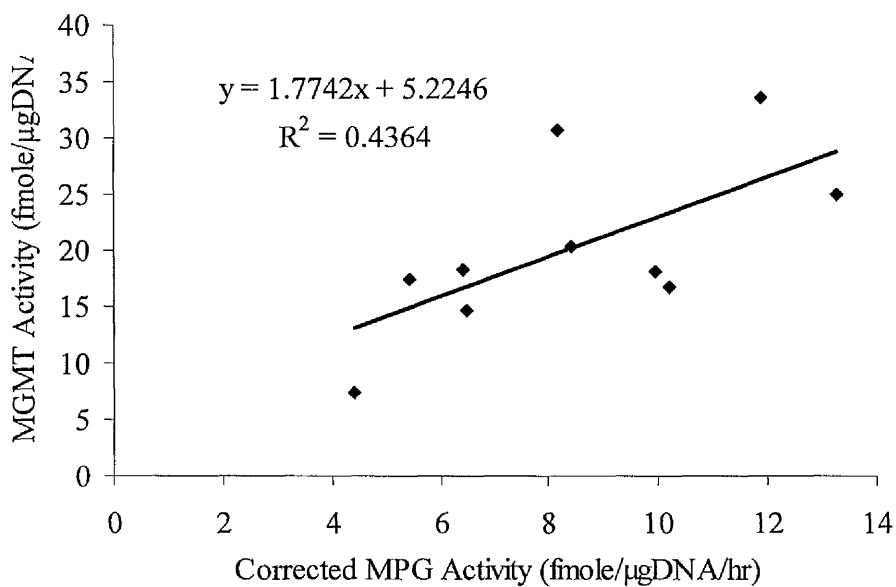
Figure 5.21a+b MGMT activity in whole PBMC samples plotted against (a) MPG and (b) corrected MPG DNA repair activity.

5.2.5 PBMC Nuclear Fraction Separation Method

To investigate the consistency of the process of nuclear fraction separation and compare repair activity in nuclear and whole cell-extract, three MCF-7 pellets were suspended in PBS, split into two equal aliquots and frozen (-80°C). The aliquots were then assayed as either a nuclear or whole cell-extract. There was a 6% difference in DNA concentration between the aliquots which was not judged significant and a four fold difference in protein concentration. Measured MPG repair activity in the nuclear extract represented 52.9% of uncorrected whole-cell extract MPG activity and 5.4% of all whole-cell extract MGMT activity (Table 5.5). The proportion of MGMT activity in the MCF-7 nuclear fraction was consistent with the findings in PBMCs (Section 5.1.4).

Table 5.5 Comparison of MCF-7 nuclear and whole sample extracts: MGMT / MPG activity, DNA and protein concentration.

Measure		MCF-7 Extracts (n = 3)	
		Nuclear	**Whole**
DNA	Mean Concentration (μg/ml) ± sd	22.0 ± 0.01	20.7 ± 0.01
Protein	Mean Concentration (mg/ml) ± sd	0.1 ± 0.0	0.4 ± 0.0
MGMT	Mean Activity (fmole/μgDNA) ± sd	0.9 ± 0.1	16.7 ± 1.6
	% of whole sample (range)	5.4 (5.3–5.7)	100
MPG	Mean Activity (fmole/μgDNA/hr) ± sd	8.3 ± 0.6	15.7 ± 0.6
	% of whole sample (range)	52.9 (43.6–67.4)	100

5.3   Discussion

This chapter describes the development and validation of oligonucleotide cleavage assays for measurement of the DNA repair proteins MGMT and MPG. At the end of this process, the MGMT and MPG assays were judged to be sensitive and accurate enough for measurement of DNA repair activity in PBMC nuclear samples (Chapter 6). Internal standards were also introduced to allow verification of consistency between assays.

Assay development resulted in several modifications to the previously used laboratory methods. Pre-blocking the 96 well microtitre plates with BSA-PBS resulted in significant improvements in the consistency and amount of substrate bound. It was demonstrated that 0.5% to 5% BSA-PBS resulted in similar improvements, 0.5% was therefore chosen. It was postulated that BSA would block non-specific binding of oligonucleotides to wells and so increase the amount of oligonucleotide specifically bound via the streptavidin-biotin bond. In non-blocked wells, non-specifically bound oligonucleotide (i.e. not bound to streptavidin) would be variably removed by washing thereby increasing inter-well variation and reducing overall binding. The introduction of pre-blocking contributed to a reduction in inter and intra-assay variability and increased the level of substrate binding overall.

The screening of a cell line was important to contribute both to the development process and for the introduction of a quality control for both assays. MCF-7 was available and expressed both MGMT and MPG cleavage activity in measurable quantities. The use of a cellular control for the assays was thought to be more appropriate than a pure enzyme control. The MCF-7 cell pellet would be processed in the same way as samples prior to each assay, so controlling for sample preparation, sonication and the cleavage assay itself.

MGMT activity was screened in a number of PBMC samples. The activity of MGMT was noted to be approximately 10 fold lower than had been reported in previous studies from the same laboratory (Margison, Heighway et al. 2005). The process of PBMC sample storage and preparation was investigated after significantly higher MGMT activity was discovered in the discarded supernatant compared to the pellet. The collected pellet was found to contain on average 98.6% of whole sample DNA and only 12.4% of sample protein. It was therefore postulated that the method of PBMC storage and processing had resulted in cellular but not nuclear lysis; possibly through the action of rapid thawing of

cells from suspension in PBS. The collection of at least 95% of sample DNA in all samples analysed suggested the method was consistent in creating a 'nuclear extract'.

The finding of 95% of MGMT activity in the 'cytoplasm' of both PBMCs and MCF-7 cells was contrary to most studies of MGMT cellular localisation using immunohistochemical (IHC) techniques, which report MGMT localisation to occur mainly in the nucleus (Ayi, Loh et al. 1992; Lee, Rafferty et al. 1992; Brent, von Wronski et al. 1993; Lim and Li 1996). However, the finding was consistent with IHC data from Ishibashi *et al*, who reported greater than 90% of MGMT to be found in cytoplasmic cellular fractions from both a human and a mouse cell line (Ishibashi, Nakabeppu et al. 1994). Belanich *et al* also reported nuclear and cytoplasmic MGMT IHC staining (Belanich, Randall et al. 1996). No previous study has examined human MGMT activity in subcellular fractions.

One reason for inconsistencies in studies of MGMT cellular localisation by IHC techniques maybe differences in the specificity of antibodies used for the active MGMT protein. Binding to inactive MGMT which is likely to be localised in the nucleus would bias results towards nuclear localisation. ELISA techniques could be used to determine the relative amounts of active and inactive protein that make up total MGMT levels in cellular fractions. However, this was outside the scope of this study. The advantage of the oligonucleotide cleavage assay over IHC staining is that only active protein is measured, so removing the possible confounding effect of inactive MGMT. Biologically it would appear sensible to have large reserves of a DNA repair protein that is constitutively expressed with no evidence of inducibility (Pieper 1997).

Mg^{2+} was added to the cleavage buffer to activate APE-1 (Erzberger and Wilson 1999). APE-1 incises the phosphodiester bond 5' to an abasic lesion creating a single strand break with a 3' hydroxyl and 5' deoxyribose phosphate end. In excess of 95% of double stranded AP site incision activity in mammals is due to APE-1 activity (Demple, Herman et al. 1991). The presence of Mg^{2+} is essential for APE-1 phosphodiester backbone cleavage, it also acts to destabilise the APE-1/incised DNA complex to facilitate APE-1 turnover (Masuda, Bennett et al. 1998). The introduction of Mg^{2+} to the cleavage buffer resulted in a marked increase in MPG substrate and control oligonucleotide cleavage by MCF-7 and PBMC nuclear and whole cell extracts. This effect was related to Mg^{2+} concentration with the greatest effect seen above 5mM and reaching a plateau by 10mM. Wilson showed that

APE-1 incision of AP sites in dsDNA was most robust at 10mM $MgCl_2$, K^+ 50mM and pH 7.5 (Wilson 2005).

A twenty fold increase in MPG substrate cleavage was seen when pure MPG was assayed with activated APE-1 compared to pure MPG without APE-1. There was no significant cleavage of the oligonucleotide by activated APE-1 alone suggesting that the increased cleavage was due to the combination of both proteins. The mechanism by which MPG substrate cleavage was increased was thought to be due to APE-1 facilitated cleavage of the phosphodiester bond 5′ to the AP site created by the removal of εA by MPG. This would significantly increase the release of the radiolabelled 5mer compared to the standard assay conditions in which APE-1 was inactivated. Studies have shown APE-1 related increased cleavage activity of the monofunctional DNA glycosylase thymine DNA glycosylase and the bifunctional DNA glycosylase hOGG1 (Waters, Gallinari et al. 1999; Cappelli, Hazra et al. 2001; Hill, Hazra et al. 2001; Vidal, Hickson et al. 2001). Xia *et al*, has also shown APE-1 associated increase in hypoxanthine base cleavage by MPG (Xia, Zheng et al. 2005). One possible mechanism to explain this effect is thought to be competitive binding of AP sites preventing the re-association of the DNA glycosylase (Vidal, Hickson et al. 2001). However, recent work by Maher *et al* showed no direct interaction between MPG and APE-1, with APE-1 not affecting the base excision activity of MPG. There was also no evidence of inhibition of MPG cleavage of an εA substrate by increasing AP sites (Maher, Vallur et al. 2007).

Adhikari *et al*, have recently shown that Mg^{2+} inhibits MPG cleavage of εA substrates; with a Mg^{2+} concentration of 5mM significantly inhibiting MPG cleavage (Adhikari, Toretsky et al. 2006). However the addition of 5mM EDTA to the same reaction effectively restored full MPG activity. This is consistent with reports of MPG inhibition by other metals including cadmium (Cd^{2+}), nickel (Ni^{2+}) and zinc (Zn^{2+}) (Wang, Guliaev et al. 2006) and reports of hOGG1 AP lyase inhibition by Mg^{2+} (Morland, Luna et al. 2005). The mechanism of inhibition of MPG by Mg^{2+} is thought to be due to a reduction in substrate binding. The use of Mg^{2+} in this study has balanced the control of repair by MPG against the regulation of AP site production. Further work would be required to clarify the exact influence of Mg^{2+} concentration in the particular parameters of this assay.

The addition of APE-1 at various concentrations to MCF-7 extract did not affect the degree of substrate cleavage. This finding was as expected; the rate limiting reaction is thought to be glycosylase cleavage. The APE-1 enzyme is found in excess in cells compared to DNA glycosylases and APE-1 phosphodiester bond 5' incision activity is 3 orders of magnitude faster than glycosylase cleavage (Strauss, Beard et al. 1997; Maher, Vallur et al. 2007).

The measurement of released radioactivity by cleavage of double stranded oligonucleotide was used to determine DNA repair activity. It was important to separate the background level of non-specific loss into the microtitre well from specific cleavage. This was achieved in the assay of each sample by including a well containing IBSA only. Comparison of visualised products using MPG substrate and pure MPG / APE-1 revealed that activity was higher in microtitre wells containing enzyme than the negative control because of the presence of a cleaved 5mer. This was consistent with cleavage of the oligonucleotide at the εA adduct. A similar finding was seen with PBMC nuclear extracts. Detected radioactivity increased with increasing concentration of extract because of increasing amount of adduct cleaved. The presence of the εA adduct was shown to be essential for substrate oligonucleotide cleavage when PBMC nuclear samples were assayed; no cleavage of a control oligonucleotide was seen. A consistent background level of non-specific release of radioactivity was seen using different nuclear PBMC samples that represented approximately 10% of MPG substrate cleavage; this was thought to be due to loss of double and single stranded oligonucleotide into the reaction well mix.

However, this was not as clear when assay products from whole PBMC cellular extracts were analysed. A single band corresponding to a 5mer product was present on both substrate and control assays. The cleavage of a 5mer product by whole PBMC extract from the MPG control oligonucleotide was a surprising finding and one that was not easily explained. Non-specific digestion of the oligonucleotide by exo or endonucleases should have produced a range of cleavage products and not a specific band. It was unlikely to be due to contamination of the control with εA containing substrate, as it occurred independently on two separately synthesised control oligonucleotides. MPG has been reported to be able to cleave normal purine based nucleotides (Berdal, Johansen et al. 1998) but there was no evidence of this when pure MPG / APE-1 was incubated with the control oligonucleotide. The lack of cleavage when the MGMT control was used could

suggest a sequence specific problem with the MPG oligonucleotide. Further studies using different MPG substrate sequences, different shaped oligonucleotide e.g. hairpin or positioning the adduct in different positions may have produced less non-specific cleavage; however this work was outside the scope of this study.

It was not clear, because the size of fragment produced was a 5mer for both the MPG substrate and control whether the process that was occurring in the control was taking place in the substrate and if so to what extent did it account for the visualised product. The cleavage of control oligonucleotide in whole PBMC extract was approximately half that seen with the substrate oligonucleotide, varying from 43 to 71%. MPG and corrected MPG activity were significantly correlated with each other and also with MGMT activity; the correlation was strongest for whole MPG activity. One previous study of PBMC MGMT and MPG activity showed no correlation between the activity of the two repair proteins (Hall, Bresil et al. 1993). However, the simple subtraction of control from substrate counts was not necessarily thought to be an accurate or appropriate correction.

Due to the practical difficulty of splitting PMBC samples, the sample preparation and processing method was validated using aliquots of MCF-7 cells. The MCF-7 nuclear pellets had equivalent DNA concentrations to the whole samples, there was also a high degree of consistency in the DNA quantification between the three nuclear sample aliquots (CV < 1%). This suggested a high proportion of nuclei were collected without lysis.

The activity of MGMT was measured in both nuclear and whole cellular extracts. 'Nuclear' MGMT activity was approximately 5% of whole sample activity for both PBMCs and MCF-7 cells. There was a high degree of consistency between values from different MCF-7 aliquots; the CV of nuclear MGMT activity was 2%. This suggested that cellular lysis was effective and consistent. Minimal variation in cell lysis would have induced a significant variation in activity due to the presence of 95% MGMT activity in the 'cytoplasm'. The proportional split of MGMT activity between nuclear and cytoplasmic fractions was highly consistent between MCF-7 and PBMC extracts.

The measurement of MPG activity in the MCF-7 cell extracts showed approximately 50% of cleavage activity was present in the nuclear sample. However, no control oligonucleotide was used to correct for non-specific cleavage in the whole cell sample. If a correlation was taken from the degree of non-specific cleavage seen in whole PBMC

extracts then the fraction of MPG activity seen in the nucleus could actually be much higher. The conclusion from the PBMC and MCF-7 experiments was that the initial method of sample processing had created a consistent subdivision of the whole cellular sample, into a pellet containing almost all cellular DNA (nuclear extract) and a supernatant consisting of 90% of cellular protein (cytoplasmic extract). It was judged appropriate to measure MGMT and MPG activity in nuclear PBMC extracts. The repair of DNA is a nuclear based process and the measurement of MGMT and MPG activity in this specific cellular fraction was therefore thought to be more relevant than whole sample DNA repair activity.

Chapter Six

MGMT and MPG Activity in PBMC
Nuclear Extracts

6.1 A Study of MGMT and MPG Repair Activity in Nuclear PBMC Extracts.

The aim of this study was to investigate whether inter-individual differences in repair activity of MGMT and MPG modify susceptibility to lung cancer. In a case-control study design, the activity of both DNA repair proteins was measured in nuclear PBMC extracts and the results compared between lung cancer cases and bronchoscopy controls selected from the main study population.

6.1.1 PBMC Study Population

Consecutive samples accounting for approximately half of the whole study population were chosen as a representative sample for DNA repair protein activity analysis (n = 243 / 471, 51.6%); referred to as the PBMC study population. A comparison of this population to the whole study revealed no significant differences in age, gender, alcohol intake, smoking exposure or the proportion of cases (Table 6.1); this suggested that the sample was representative of the study as a whole.

Table 6.1 A comparison of the PBMC and whole study populations.

Variables	Whole Study n = 471	PBMC Study n = 243	p value
Cases / Controls (% cases)	164 / 307 (34.8)	83 / 160 (34.2)	0.76
Age mean (yrs) ± sd	67.0 ± 10.5	66.4 ± 11.0	0.21
Sex M / F (% male)	297 / 174 (63.1)	147 / 96 (60.5)	0.23
Alcohol mean (units/wk) ± sd	15.6 ± 34.9	13.9 ± 28.7	0.19
Ever Smoked n = yes (% yes)	441 (93.6)	224 (92.2)	0.18
Current Smokers n = yes (%[b])	168 (38.1)	80 (35.7)	0.20
Smoking Duration[a] mean (yrs) ± sd	42.1 ± 13.4	41.3 ± 13.7	0.20
Packyears[a] mean ± sd	52.0 ± 34.0	51.1 ± 34.0	0.61
Age Started Smoking[a] mean (yrs) ± sd	16.5 ± 4.3	16.4 ± 4.0	0.51
Cigarettes / day[a] mean ± sd	24.0 ± 13.8	24.1 ± 14.2	0.96

[a]in ever smokers only

6.1.2 MGMT Study

A total of 243 samples were assayed for MGMT activity and 91.8% (n = 223 / 243) of these results were accepted. Samples were excluded if the MCF-7 cellular control value of the assay fell outside the accepted range (Section 5.1.9.1). Those subjects with accepted MGMT results were referred to as the MGMT study population. A comparison of this group to the whole study population revealed no significant differences with regard to age, sex, proportion of cases, alcohol or smoking exposure (Table 6.2). Accepted samples were therefore judged to be representative of the study as a whole.

Table 6.2 Comparison of the whole study and MGMT study populations.

Variable	Whole Study n = 471	MGMT Study n = 223	p value
Cases / Controls (% cases)	164 / 307 (34.8)	74 / 149 (33.2)	0.60
Age mean (yrs) ± sd	67.0 ± 10.5	66.3 ± 11.1	0.21
Sex M / F (% male)	297 / 174 (63.1)	133 / 90 (59.6)	0.18
Alcohol mean (units/wk) ± sd	15.6 ± 34.9	15.4 ± 35.9	0.24
Ever Smoked n = yes (% yes)	441 (93.6)	206 (92.4)	0.22
Current Smokers n = yes (%[b])	168 (38.1)	73 (32.7)	0.19
Smoking Duration[a] mean (yrs) ± sd	42.1 ± 13.4	41.0 ± 13.6	0.13
Packyears[a] mean ± sd	52.0 ± 34.0	52.3 ± 34.7	0.87
Age Started Smoking[a] mean (yrs) ± sd	16.5 ± 4.3	16.5 ± 4.1	0.68
Cigarettes / day[a] mean ± sd	24.0 ± 13.8	24.6 ± 14.4	0.66

[a] in ever smokers only

MGMT activity was detected in 97.8% of samples (n = 218 / 223). Values ranged from 0.29 - 13.37 fmole/μgDNA, this was a 46.1 fold variation. The five samples with activity below the limit of quantitation of the assay were given values half that of the limit of quantitation for the purpose of statistical analysis (0.11 fmole/μgDNA). MGMT activity in the whole population was not normally distributed (Figure 6.1), therefore MGMT data has been presented as geometric means ± standard deviations.
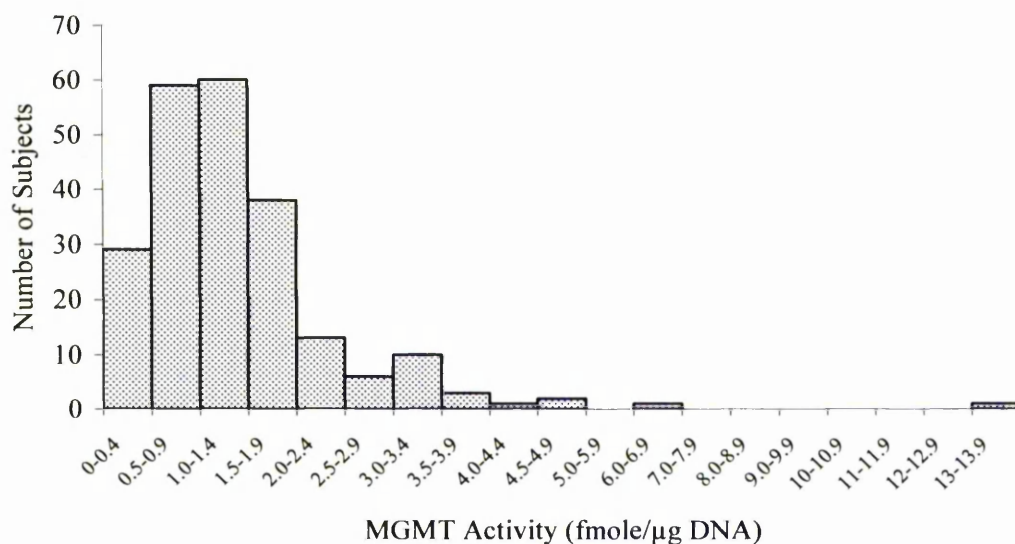
Figure 6.1 Distribution of MGMT activity.

6.1.2.1 MGMT Activity and Case Status

The MGMT study population consisted of 74 cases and 149 controls (n = 223). A comparison of cases and controls is detailed in table 6.3. Smoking duration in ever smokers was the only significant difference, with cases having smoked for an average of five years longer than controls (44.2 ± 12.0 vs. 39.3 ± 14.1 years; $p = 0.02$). Cases were also more likely to be current smokers (40.5% vs. 28.9%, $p = 0.08$), a difference that approached significance.

The distribution of MGMT activity in cases and controls is shown in figure 6.2. The geometric mean of MGMT activity in cases was not significantly different to that in controls (1.15 ± 1.98 vs. 1.10 ± 2.07 fmole/μgDNA, $p = 0.65$) (Table 6.3). There was also no difference between cases and controls after cases were stratified according to histological subtype (Table 6.4).

Table 6.3 MGMT study population, comparison of cases and controls.

Variable	Cases n = 74	Controls n = 149	p value
MGMT Activity (fmole/µgDNA) geometric mean ± sd	1.10 ± 2.07	1.15 ± 1.98	0.65
Age mean (yrs) ± sd	67.4 ± 10.7	65.8 ± 11.2	0.32
Sex M / F (% male)	49 / 25 (66.2)	84 / 65 (56.4)	0.16
Alcohol mean (units/wk) ± sd	19.5 ± 48.1	13.4 ± 27.9	0.31
Ever Smoked n = yes (% yes)	71 / 3 (95.9)	135 / 14 (90.6)	0.16
Current Smokers n = yes (%[b])	30 (40.5)	43 (28.9)	0.08
Smoking Duration[a] mean (yrs) ± sd	44.2 ± 12.0	39.3 ± 14.1	**0.02**
Packyears[a] mean ± sd	53.1 ± 29.9	51.8 ± 37.0	0.80
Age Started Smoking[a] mean (yrs) ± sd	15.9 ± 3.5	16.7 ± 4.4	0.17
Cigarettes / day[a] mean ± sd	23.7 ± 11.8	25.1 ± 15.6	0.49

[a]in ever smokers only
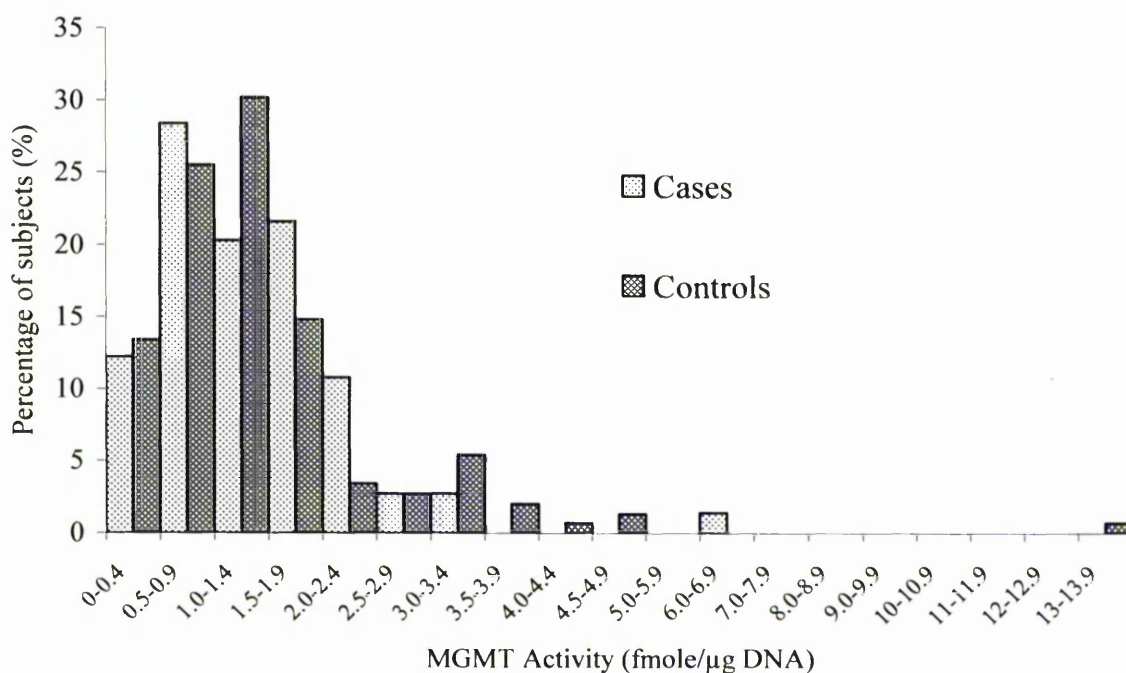
Figure 6.2 Distribution of MGMT activity in cases and controls.

Table 6.4 MGMT activity in cases compared to controls

after stratification for histological type.

Lung Cancer Histology	Number of Subjects (n)	MGMT Activity (fmole/µgDNA) Geometric mean ± sd	p value[a]
Controls	149	1.10 ± 2.07	-
Squamous Cell	20	1.20 ± 1.97	0.61
Adenocarcinoma	19	1.20 ± 1.74	0.61
Small Cell	13	1.12 ± 1.99	0.35
Large Cell	10	1.03 ± 2.15	0.77
No Histology	10	0.87 ± 2.46	0.33

[a]compared with controls

MGMT activity did increase with advancing lung cancer stage in subjects with NSCLC, but this was a non-significant trend (p = 0.69) (Table 6.5). This analysis was limited to subjects diagnosed with NSCLC as all cases with SCLC had been diagnosed with advanced disease.

Table 6.5 MGMT activity in NSCLC cases stratified according to stage.

Lung Cancer Stage	Number of Subjects	MGMT Activity (fmole/µgDNA) Geometric mean ± sd	p value
Early (Tx-1B)	6	0.99 ± 1.70	
Intermediate (2A-3A)	9	1.12 ± 1.72	0.69
Advanced (3B-4)	32	1.20 ± 2.03	

6.1.2.2 MGMT Activity and Age

MGMT activity was not related to age when analysed in the whole MGMT study population ($R^2 = 0.001$, $p = 0.67$) (Figure 6.3). MGMT activity did not vary according to case status either above ($p = 0.35$) or below ($p = 0.12$) the median age of controls. There was also no difference in MGMT repair activity between cases in either age category ($p = 0.50$) (Table 6.6).

The equation and statistics shown on the figure:
$$y = 0.0009x - 0.0134$$
$$R^2 = 0.0012$$
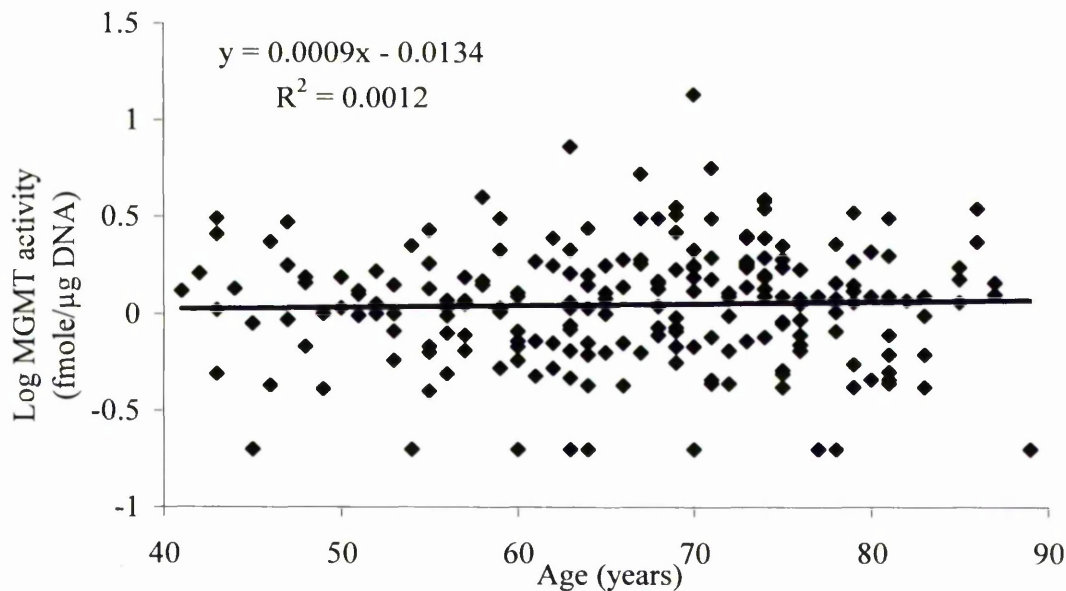
Figure 6.3 log MGMT activity vs. age in the whole MGMT study.

Table 6.6 MGMT activity in cases and controls stratified for median age of controls.

Age Category[a] (years)	Number of subjects cases / controls	MGMT Activity (fmole/µg DNA) geometric mean ± sd		p value
		Cases	**Controls**	
≤ 67.0	33 / 77	1.23 ± 2.07	0.98 ± 1.98	0.12
> 67.0	41 / 72	1.10 ± 1.92	1.25 ± 2.13	0.35
p value		0.50		

[a]based on the median age of controls

6.1.2.3 MGMT Activity and Gender

There was no significant difference in MGMT activity according to gender in the MGMT study population (male 1.16 ± 2.03 vs. female 1.06 ± 2.05 fmole/µgDNA; $p = 0.35$). MGMT activity did not vary according to case status after stratification for gender (Table 6.7). There was also no significant difference between MGMT activity in male and female cases ($p = 0.79$).

Table 6.7 MGMT activity in cases and controls stratified for gender.

Gender	Number of subjects cases / controls	MGMT Activity (fmole/µg DNA) geometric mean ± sd		p value
		Cases	**Controls**	
Male	49 / 84	1.14 ± 2.07	1.17 ± 2.02	0.79
Female	25 / 65	1.19 ± 1.83	1.01 ± 2.14	0.35
p value		0.79		

6.1.2.4 MGMT Activity and Smoking

In the MGMT population as a whole, there was no significant difference in MGMT activity between current ($p = 0.41$) or former smokers (0.30) and never smokers (Table 6.8).

Table 6.8 Smoking status and MGMT activity.

Smoking Status	Number of subjects	MGMT Activity (fmole/µg DNA) geometric mean ± sd	p value[a]
Never	17	0.95 ± 1.68	-
Current	73	1.10 ± 2.01	0.41
Former	133	1.15 ± 2.10	0.30

[a]compared to never smokers

MGMT activity in ever smoking subjects was not related to time since last cigarette (Figure 6.4), smoking duration (p = 0.70), packyears (p = 0.30), or age smoking started (p = 0.94) (Figure 6.5a-c).
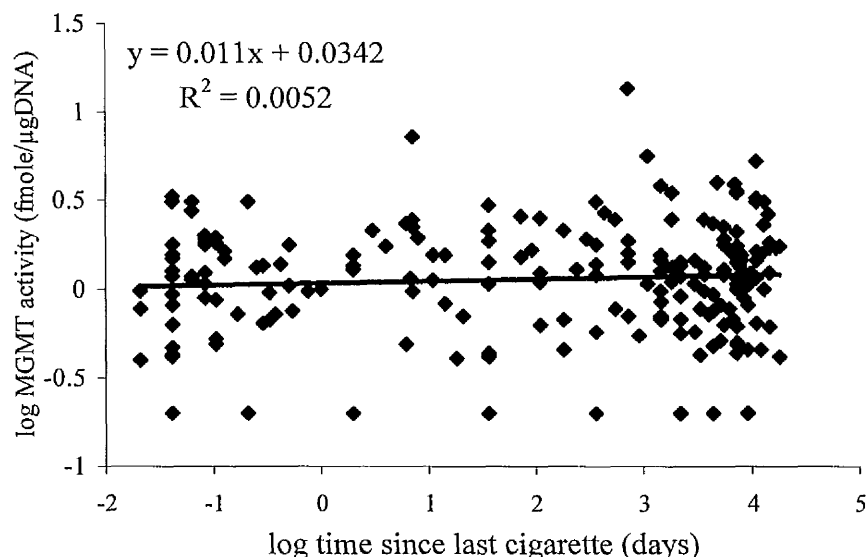
Figure 6.4 log MGMT activity against the log of time since last cigarette (days).

Subjects were stratified according to case and smoking status. Current smoking cases had a significantly higher MGMT repair activity than controls (1.40 ± 1.95 vs. 0.93 ± 1.96 fmole/μg DNA; p = 0.01) (Table 6.9). There was no difference according to case status in never (p = 0.85) or former smokers (p = 0.18), although with only three never smoking cases this comparison was very limited. In cases, MGMT activity was significantly greater in current smokers when compared to former smokers (p = 0.05).

Table 6.9 MGMT activity in cases and controls after stratifying for smoking status

Smoking Status	Number of Subjects (cases / controls)	MGMT Activity (fmole/μg DNA) geometric mean ± sd		p value
		Cases	Controls	
Never	3 / 14	1.00 ± 2.11	0.94 ± 1.64	0.85
Current	30 / 43	1.40 ± 1.95	0.93 ± 1.96	**0.01**
Former	41 / 92	1.01 ± 1.96	1.22 ± 2.16	0.18

(a) Smoking duration

$y = -0.0006x + 0.0781$
$R^2 = 0.0006$

(Log MGMT Activity vs Smoking Duration (years))

(b) Packyears

$y = -0.0007x + 0.0888$
$R^2 = 0.0053$

(Log MGMT Activity vs Packyears)

(c) Age Started

$y = 0.0018x + 0.0254$
$R^2 = 0.0005$
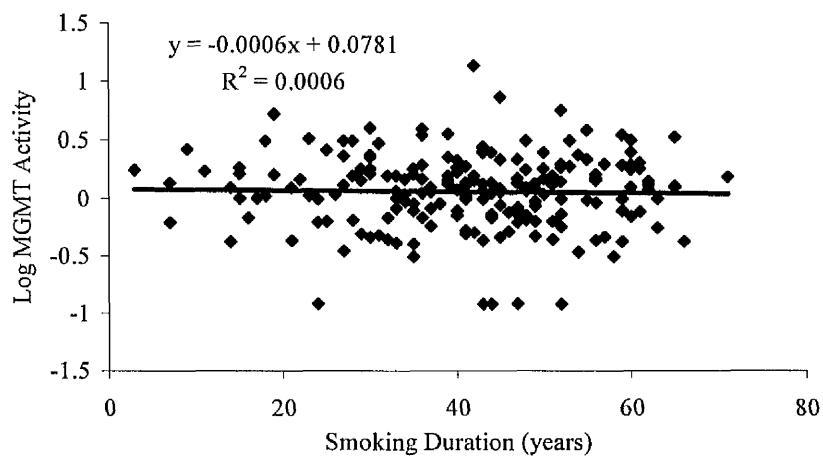
(Log MGMT Activity vs Age Started (years))

Figure 6.5a-c Log MGMT activity and (a) smoking duration, (b) packyears and (c) age started smoking in ever smokers.

6.1.2.5 MGMT Activity and Genotype

MGMT intron 1 genotyping was performed successfully on 219 / 223 (98.2%) subjects; codon 178 genotyping was determined on 222 / 223 (99.6%) of subjects. MGMT activity was analysed according to genotype in cases and controls. There was no significant association with nuclear MGMT activity for either genotype, although numbers were small for codon 178 GG genotype (Table 6.11). The lowest activity level was seen in the TT genotype for both cases and controls.

Table 6.11 MGMT intron 1 and codon 178 genotype and MGMT activity.

MGMT Genotype	Category	Case/ Control	MGMT Activity (fmole/µg DNA) geometric mean ± sd		p value
			Cases	**Controls**	
Intron 1	**CC**	38 / 71	1.25 ± 1.94	1.15 ± 2.12	0.57
	CT	25 / 61	1.12 ± 1.82 $p = 0.50^{a}$	1.09 ± 2.17	0.88
	TT	11 / 13	0.92 ± 2.51 $p = 0.07^{a}$	0.88 ± 1.53	0.86
	p value		0.42		
Codon 178	**AA**	51 / 110	1.24 ± 1.90	1.10 ± 2.04	0.32
	AG	22 / 34	0.95 ± 2.13 $p = 0.14^{b}$	1.13 ± 2.20	0.44
	GG	1 / 4	2.14 $p = 0.40^{b}$	0.72 ± 1.53	0.10
	p value		0.22		

[a]compared to CC, [b]compared to AA

6.1.3   MPG Study

A total of 243 samples were assayed for MPG activity and 57.2% (n = 139 / 243) of results were accepted. Samples were excluded if the MCF-7 cellular control value for the assay fell outside the accepted range (Section 5.1.9.1). Those subjects with accepted MPG results were referred to as the MPG study population. A comparison of the whole and MPG study populations revealed no significant differences with regard to age, sex, proportion of cases, alcohol intake or smoking exposure (Table 6.12). The MPG population was therefore judged to be representative of the study as a whole.

Table 6.12 Comparison of whole and MPG study population.

Variable	Whole Study n = 471	MPG Study n = 139	p value
Cases / Controls (% cases)	164 / 307 (34.8)	51 / 88 (36.7)	0.82
Age mean (yrs) ± sd	67.0 ± 10.5	65.6 ± 11.3	0.09
Sex M / F (% male)	297 / 174 (63.1)	82 / 57 (59.0)	0.19
Alcohol mean (units/wk) ± sd	15.6 ± 34.9	13.6 ± 30.6	0.58
Ever Smoked n = yes (% yes)	441(93.6)	128 (92.1)	0.23
Current Smokers n = yes (%[b])	168 (38.1)	49 (35.3)	0.52
Smoking Duration[a] mean (yrs) ± sd	42.1 ± 13.4	41.7 ± 14.0	0.41
Packyears[a] mean ± sd	52.0 ± 34.0	49.4 ± 30.5	0.35
Age Started Smoking[a] mean (yrs) ± sd	16.5 ± 4.3	16.1 ± 3.1	0.20
Cigarettes / day[a] mean ± sd	24.0 ± 13.8	23.2 ± 12.4	0.61

[a]in ever smokers only
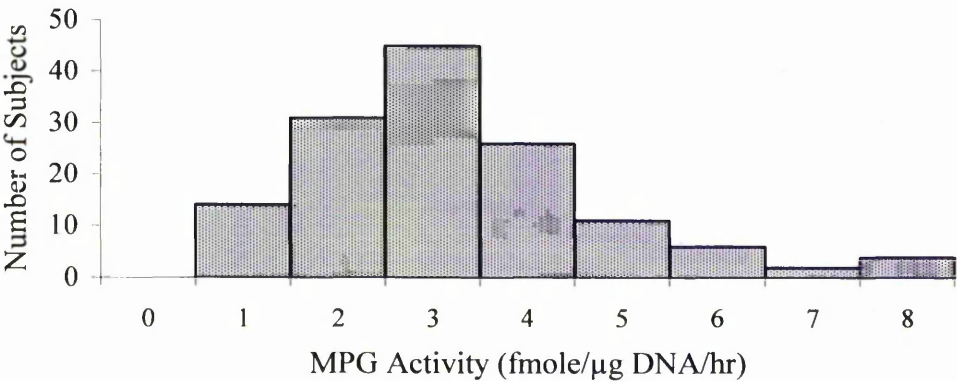
Figure 6.6 The distribution of MPG activity in the study population.

MPG repair activity was detected in all samples (n = 139), with values ranging from 1.37 – 8.43 fmole/μgDNA/hour; this was a 6.2 fold variation. MPG had a normal distribution (Figure 6.6); MPG data has therefore been presented as means ± standard deviations.

6.1.3.1 MPG Activity and Case Status

The MPG study population consisted of 88 controls and 51 cases (n = 139). A comparison of cases and controls is detailed below (Table 6.13). Differences between the groups that approached significance included a greater proportion of males in cases (68.6% vs 53.4%, p = 0.08) and in ever smoking subjects, the duration of smoking was on average five years longer in cases than controls (44.8 ± 12.0 vs. 39.9 ± 14.9 years, p = 0.06).

The mean MPG activity in cases was 4.21 ± 1.67 fmole/μgDNA/hr, this was significantly greater than the mean activity in controls (p = 0.005), which was 3.47 ± 1.35 fmole/μgDNA/hr. The distribution of MPG activity after stratification for case status is shown in figure 6.7.

Table 6.13 Comparison of MPG activity, age, gender, alcohol and smoking exposure between cases and controls.

Variable	Cases n = 51	Controls n = 88	p value
MPG Activity (fmole/μg DNA/hr) mean ± sd	4.21 ± 1.67	3.47 ± 1.35	**0.005**
Age mean (yrs) ± sd	66.9 ± 11.1	64.8 ± 11.4	0.29
Sex M / F (% male)	35 / 16 (68.6)	47 / 41 (53.4)	0.08
Alcohol mean (units/wk) ± sd	17.0 ± 37.3	11.7 ± 26.1	0.33
Ever Smoked n = yes (% yes)	48 / 3 (94.1)	80 / 8 (90.9)	0.50
Current Smokers n = yes (%[b])	19 (37.3)	30 (34.1)	0.71
Smoking Duration[a] mean (yrs) ± sd	44.8 ± 12.0	39.9 ± 14.9	0.06
Packyears[a] mean ± sd	52.5 ± 28.0	47.5 ± 31.9	0.37
Age Started Smoking[a] mean (yrs) ± sd	15.8 ± 2.9	16.2 ± 3.3	0.54
Cigarettes / day[a] mean ± sd	23.4 ± 11.1	23.2 ± 13.2	0.92

[a]in ever smokers only

(a) Cases
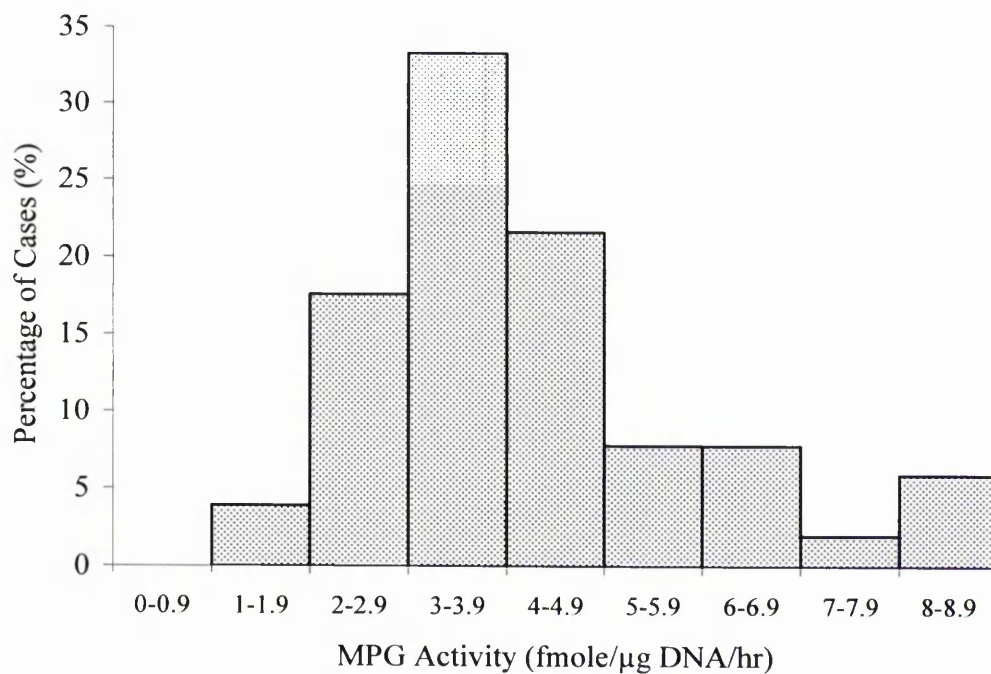
(b) Controls
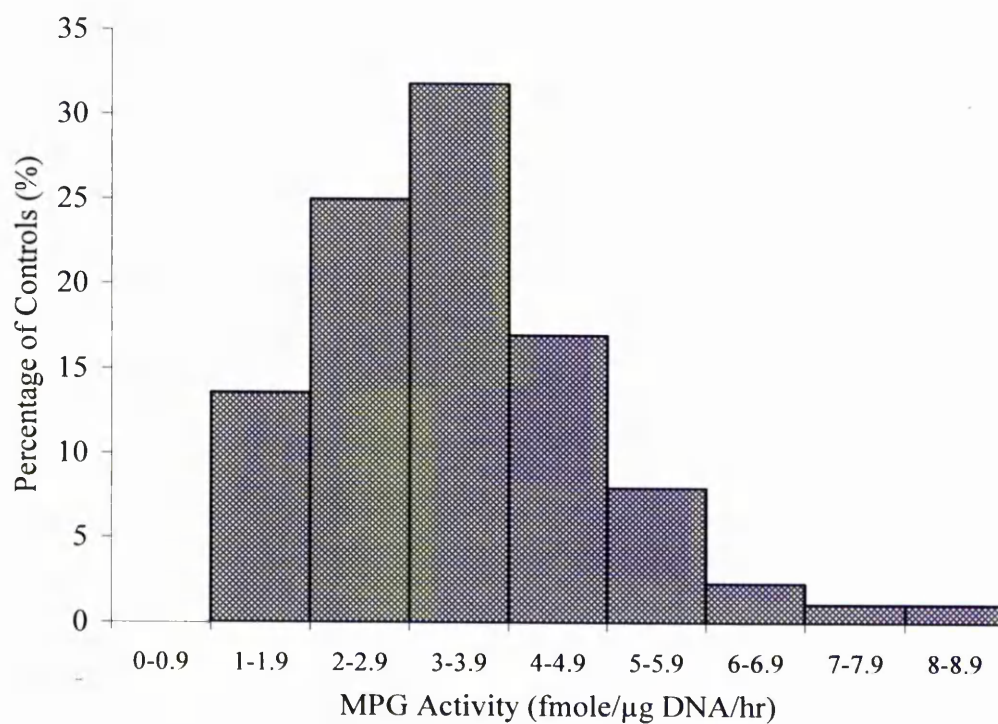
Figure 6.7 The distribution of MPG activity in (a) cases and (b) controls.

MPG activity was higher in cases with all subtypes of tumour than controls, after cases were stratified according to histology (Table 6.14). This difference reached significance in small cell carcinoma, the PBMCs of which had the highest activity of all tumour types (4.69 ± 1.66 fmole/μgDNA/hr, p = 0.01) and approached significance for adenocarcinoma (4.24 ± 1.47 fmole/μg DNA/hr, p = 0.08). The lowest activity was seen in squamous cell carcinoma (3.68 ± 1.40 fmole/μg DNA/hr). When all cases with NSCLC were grouped together, the mean MPG activity was significantly greater than controls (4.09 ± 1.74 fmole/μg DNA/hr, p = 0.04) but lower than cases with small cell carcinomas, however this did not reach significance (p = 0.25).

Table 6.14 MPG repair activity in cases stratified for histological type.

Lung Cancer Histology	Number of Subjects (n)	MPG Activity (fmole/μg DNA/hr) mean ± sd	p value[a]
Controls	88	3.47 ± 1.35	-
Squamous Cell	13	3.68 ± 1.40	0.60
Adenocarcinoma	11	4.24 ± 1.47	0.08
Small Cell	10	4.69 ± 1.66	**0.01**
Large Cell	7	4.16 ± 2.45	0.48
No Histology	7	3.82 ± 1.10	0.50

[a]compared with controls

MPG activity was not associated with stage of disease at diagnosis in subjects with NSCLC (p = 0.38) (Table 6.15).

Table 6.15 MPG activity in NSCLC cases stratified according to stage.

Lung Cancer Stage	Number of Subjects	MPG Activity (fmole/μgDNA/hr) mean ± sd	p value
Early (Tx-1B)	4	3.80 ± 1.64	
Intermediate (2A-3A)	8	3.25 ± 1.84	0.38
Advanced (3B-4)	18	4.54 ± 1.75	
Unknown	2	3.98 ± 0.97	

6.1.3.2 MPG Activity and Age

MPG activity was not related to age when analysed in the study population as a whole (R^2 < 0.01, p = 0.90) (Figure 6.8). MPG activity was also unrelated to age when subjects were stratified according to case status and to age above and below the median age of controls. MPG activity was higher in cases compared to controls in both categories, this reached significance only in subjects below the median age ($\leq$ 65.5 years) (4.24 $\pm$ 1.76 vs. 3.32 $\pm$ 1.20 fmole/μgDNA/hr, p = 0.03) (Table 6.16).

The equation on the figure reads: $y = -0.0014x + 3.8319$, $R^2 = 0.0001$.
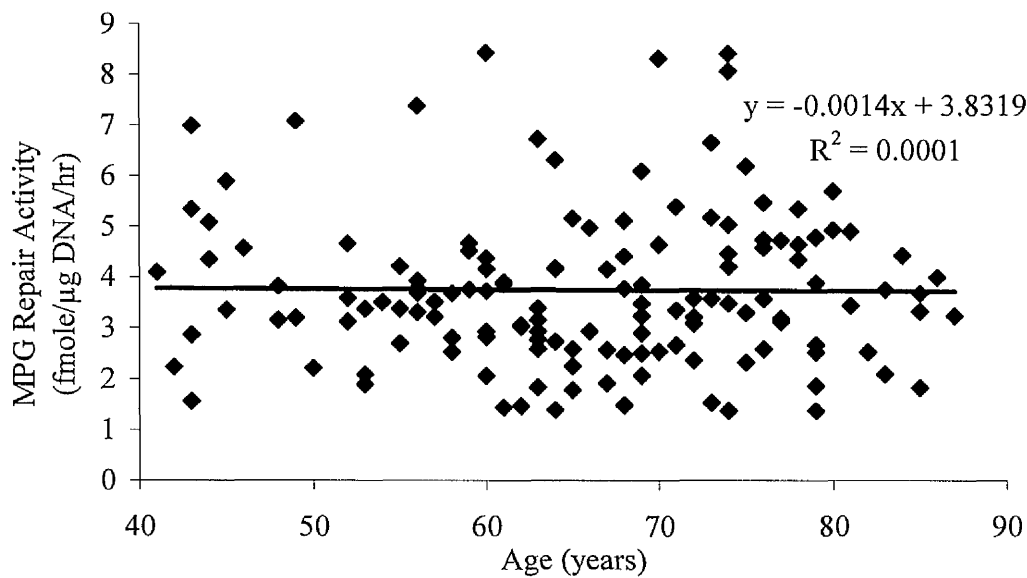
Figure 6.8 MPG activity and age in the whole study population.

Table 6.16 MPG activity in cases and controls stratified for median age of controls.

Age Category[a] (years)	Cases / Controls	MPG Activity (fmole/μgDNA/hr) mean $\pm$ sd.		p value
		Cases	Controls	
$\leq$ 65.5	24 / 44	4.24 $\pm$ 1.76	3.32 $\pm$ 1.20	**0.03**
> 65.5	27 / 44	4.19 $\pm$ 1.61	3.61 $\pm$ 1.49	0.13
p value		0.91		

[a]based on the median age of controls.

6.1.3.3 MPG Activity and Gender

There was no significant difference in MPG activity between males (3.75 ± 1.45 fmole/μgDNA/hr) and females (3.73 ± 1.62 fmole/μgDNA/hr) in the study population as a whole (p = 0.94). However, stratification according to case status and gender showed female cases to have the highest mean MPG activity of 4.93 ± 1.87 fmole/μgDNA/hr, significantly higher than both female controls (p = 0.003) and male cases (p = 0.04) (Table 6.17). The distribution of MPG activity in cases and controls according to gender is shown in figure 6.9 and 6.10.

Table 6.17 MPG activity in cases and controls according to gender.

Gender	Cases / Controls	MPG Activity (fmole/μgDNA/hr) mean ± sd.		p value
		Cases	Controls	
Male	35 / 47	3.88 ± 1.48	3.65 ± 1.43	0.47
Female	16 / 41	4.93 ± 1.87	3.26 ± 1.24	**0.003**
p value		**0.04**		
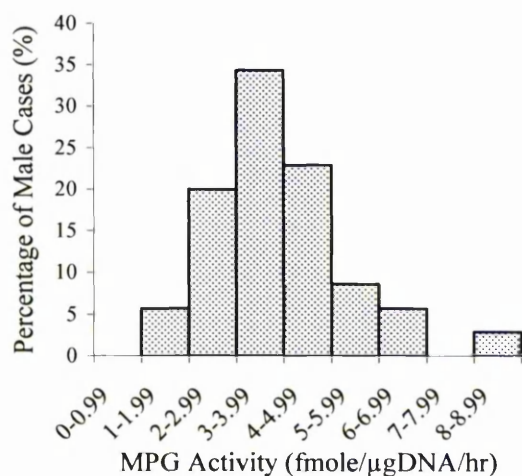
(a) Male Cases   **(b) Male Controls**
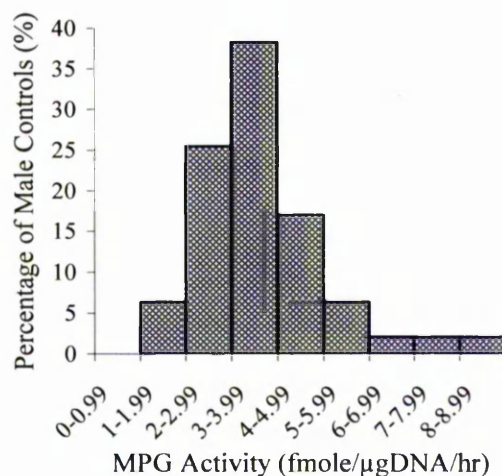
Figure 6.9 Distribution of MPG activity in male: (a) cases and (b) controls.
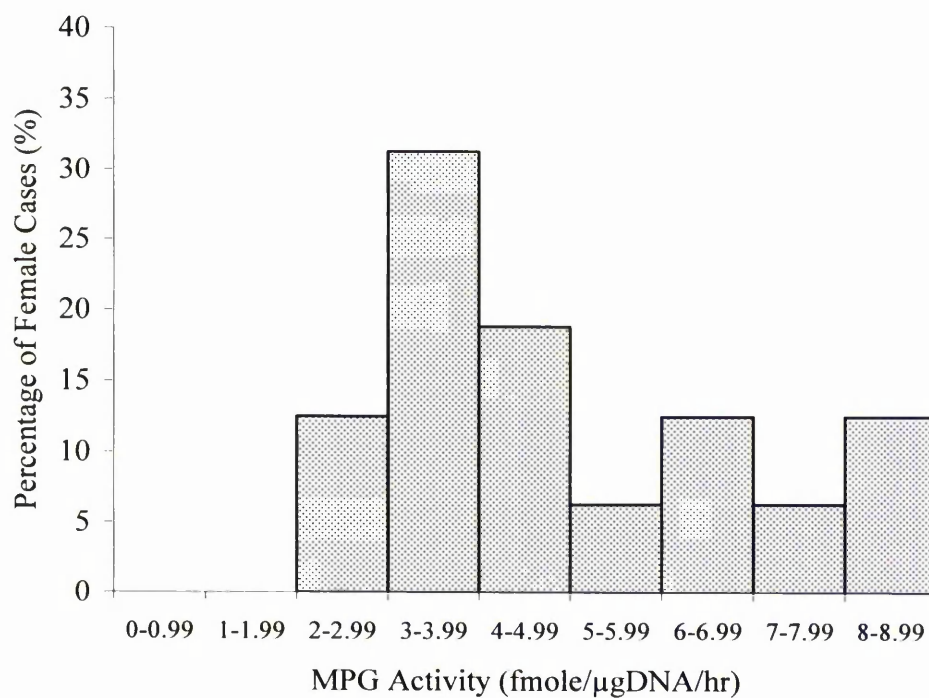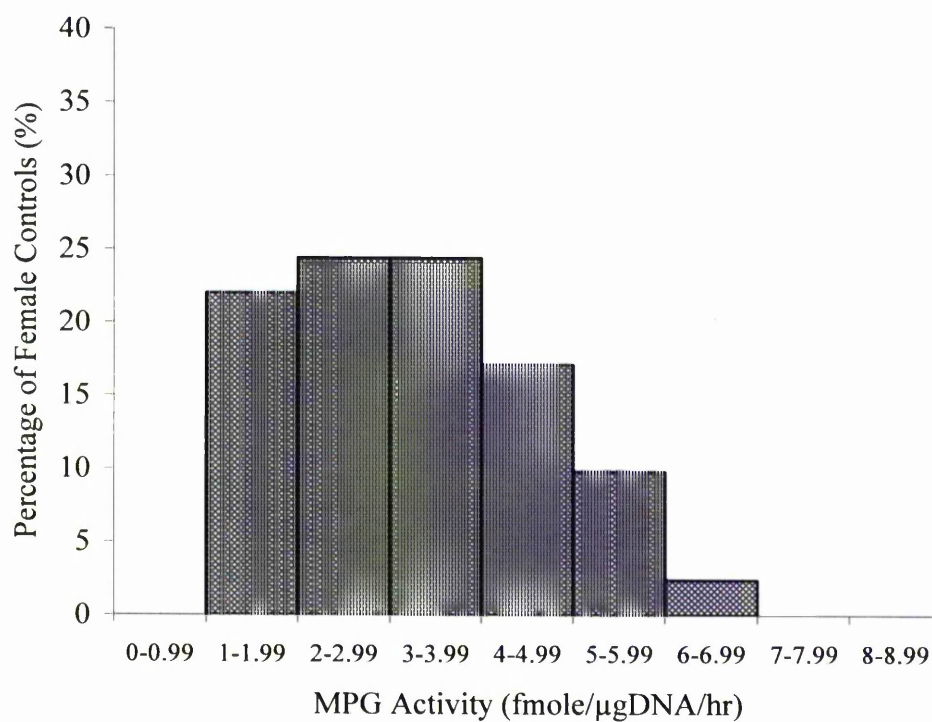
Figure 6.10 Distribution of MPG activity in female: (a) cases and (b) controls.

6.1.3.4 MPG Activity and Smoking Exposure

MPG activity was higher in ever smoking subjects (n = 128) (3.81 ± 1.53 fmole/μgDNA/hr) compared to lifelong non-smokers (n = 11) (2.95 ± 1.04 fmole/μgDNA/hr), a difference that approached significance (p = 0.07). There was no difference in MPG activity between current (n = 49) and former smokers (n = 79) (3.84 ± 1.56 vs. 3.79 ± 1.52 fmole/μgDNA/hr, p = 0.86). Both current and former smokers had greater activity than lifelong non-smokers, a difference that approached significance (p = 0.08 for both). There was no relationship between time of last cigarette and MPG activity (Figure 6.11).
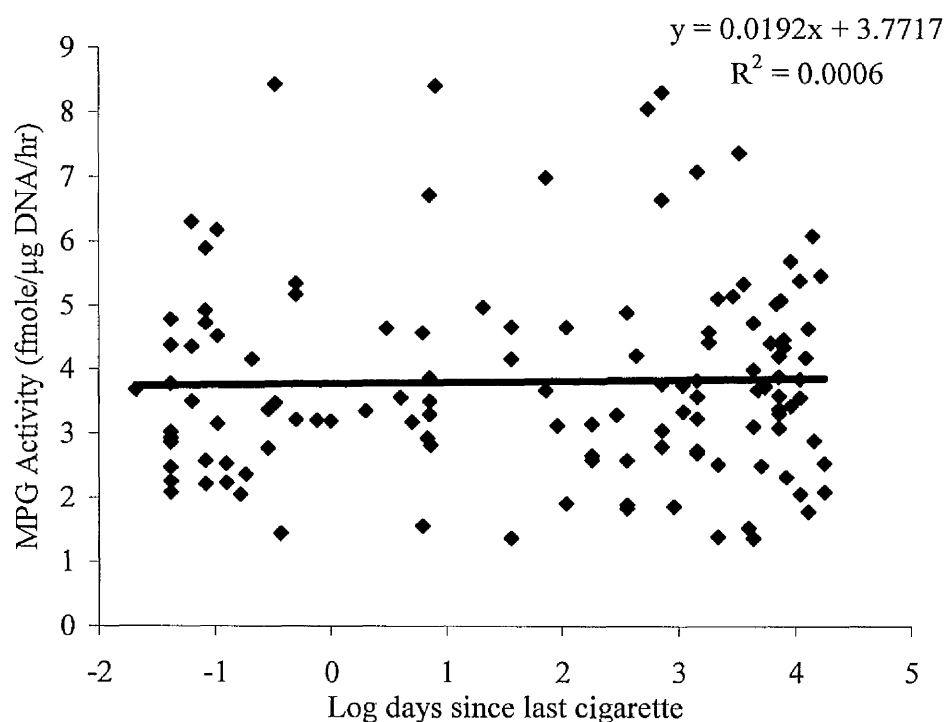
Figure 6.11 MPG activity and time since last cigarette in ever smoking subjects.

Ever smoking cases had a significantly greater MPG activity than ever smoking controls (4.26 ± 1.70 vs. 3.53 ± 1.36 fmole/μg DNA/hr, p = 0.009). After further stratification for smoking status, only current smoking cases had a significantly higher MPG repair activity than controls (4.46 ± 1.93 vs. 3.44 ± 1.14 fmole/μg DNA/hr, p = 0.05). The activity for former and never smokers was higher in cases for both categories than controls but this did not reach significance (Table 6.18).

Table 6.18 MPG repair activity after stratifying for smoking status

Smoking Status	Number of Subjects (cases / controls)	MPG Activity (fmole/μg DNA/hr) mean ± sd		p value
		Cases	Controls	
Never	3 / 8	3.43 ± 0.75	2.77 ± 1.11	0.38
Current	29 / 50	4.46 ± 1.93	3.44 ± 1.14	**0.05**
Former	19 / 30	4.13 ± 1.56	3.59 ± 1.48	0.13

The age at which a subject started to smoke did have a weak but significant association with MPG activity ($R^2 = 0.09$, $p = 0.0005$). This association was strongest in ever smoking cases ($R^2 = 0.32$, $p = 0.0002$) (Figure 6.12) and significant in ever smoking controls, but with a very low R^2 value ($R^2 = 0.03$, $p = 0.003$).

$$y = 0.3357x - 1.055$$
$$R^2 = 0.3239$$
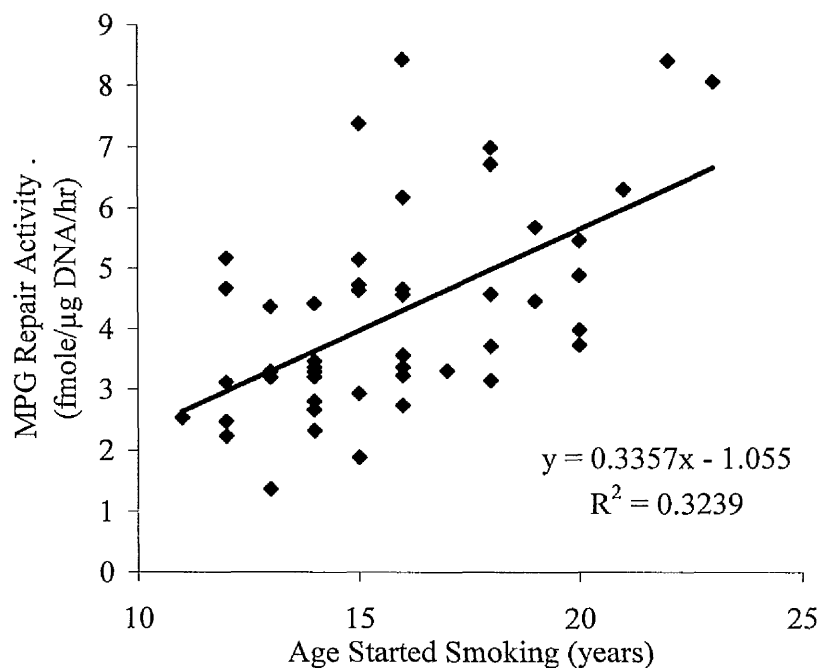
Figure 6.12 MPG activity in ever smoking cases versus age smoking started.

218

Smoking duration was not associated with MPG activity in ever smokers (p = 0.66) (Figure 6.13).

The scatter plot shows MPG Activity (fmole/µg DNA/hr) on the y-axis (0 to 9) versus Smoking Duration (years) on the x-axis (0 to 80). The regression equation is:

$$y = -0.0044x + 3.9884$$
$$R^2 = 0.0016$$
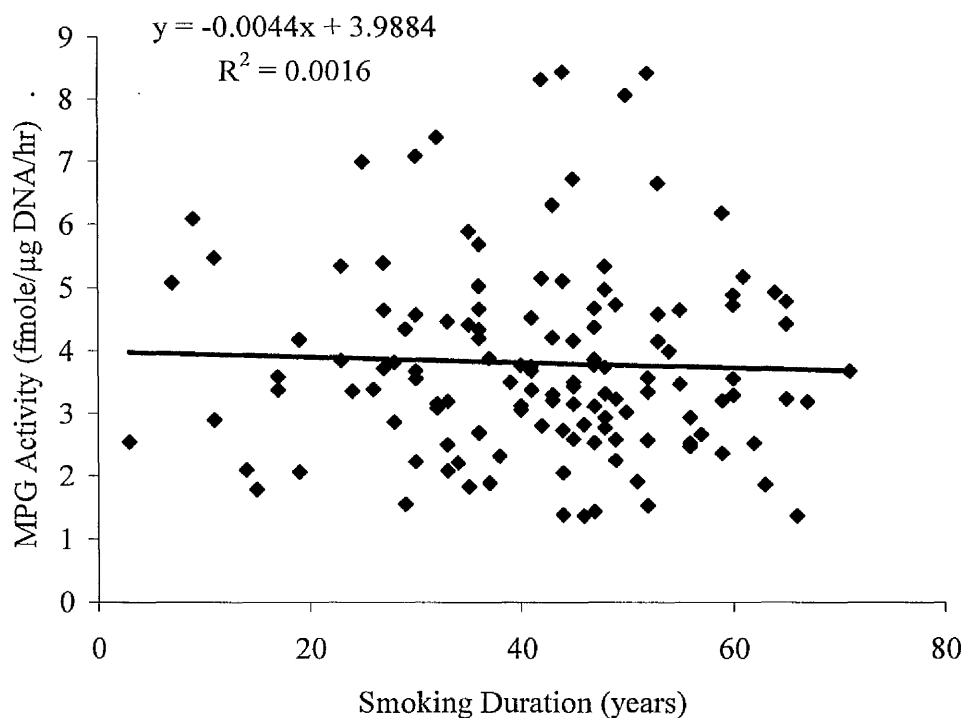
Figure 6.13 MPG activity and smoking duration in ever smokers only.

There was no difference in MPG activity between ever smoking cases or controls after stratification according to median smoking duration in controls. However cases in the above median smoking duration category had a significantly greater MPG activity than controls (4.29 ± 1.75 vs. 3.40 ± 1.46, p = 0.02) (Table 6.19).

Table 6.19 MPG activity in ever smoking cases and controls stratified for median smoking duration.

Smoking Duration Category[a] (years)	Cases/ Controls n	MPG Activity (fmole/µg DNA/hr) mean ± sd		p value
		Cases	Controls	
≤ 41.0	15 / 41	4.19 ± 1.65	3.67 ± 1.26	0.38
> 41.0	33 / 39	4.29 ± 1.75	3.40 ± 1.46	**0.02**
p value		0.84		

[a]based on the median smoking duration of ever smoking controls.

219

In ever smokers, packyear exposure was not associated with MPG activity (Figure 6.14).
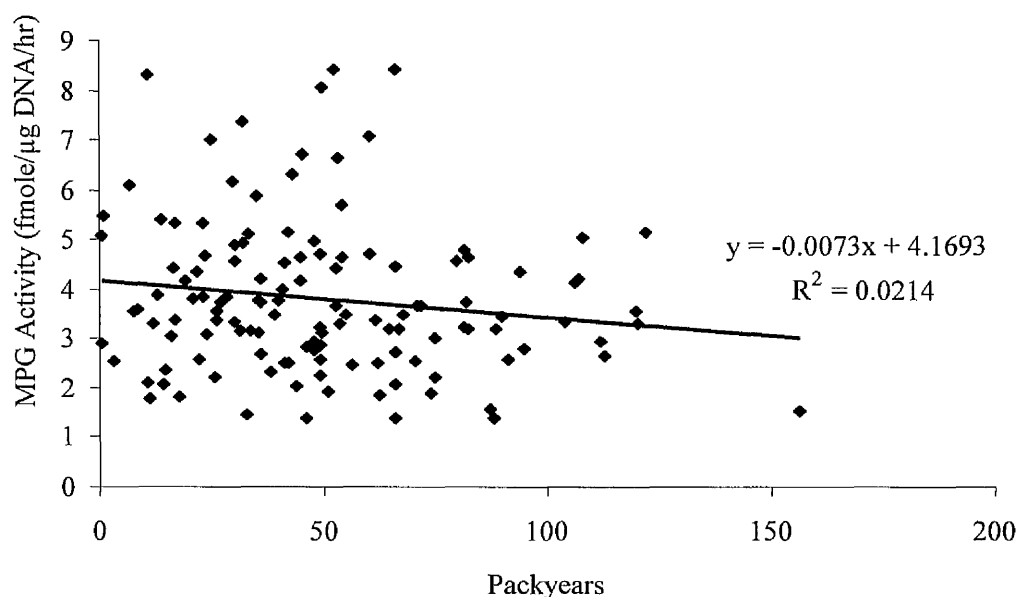
Figure 6.14 MPG activity and packyears in ever smokers only.

There was no difference in MPG activity between ever smoking cases or controls after stratification according to median number of packyears in controls. However cases in the above median packyears category had a significantly greater MPG activity than controls (4.29 ± 1.75 vs. 3.40 ± 1.46, p = 0.04), a similar finding was seen in the below median packyear category with a difference that approached significance (p = 0.09) (Table 6.20).

Table 6.20 MPG activity in ever smoking cases and controls stratified for median packyears.

Packyear Category[a]	Cases / Controls	MPG Activity (fmole/µg DNA/hr) mean ± sd.		p value
		Cases	Controls	
≤ 41.5	17 / 40	4.43 ± 1.39	3.73 ± 1.40	0.09
> 41.5	33 / 39	4.17 ± 1.86	3.34 ± 1.31	**0.04**
p value		0.62		

[a]based on the median packyears of ever smoking controls

6.1.3.5 MPG Activity and Lung Cancer Incidence

To illustrate the effect of MPG activity on the odds of lung cancer incidence a binary logistic regression was performed with case status as the outcome variable. MPG activity was entered into the calculation as a continuous variable and adjustment was made for age, gender and smoking duration (Table 6.21). The adjusted odds ratio of lung cancer incidence were significantly increased with each unit increase in MPG activity (OR_{adj} 1.41, 95% CI 1.10-1.80; p = 0.007) when analysed in the whole MPG study population.

Table 6.21 MPG activity (continuous variable) and odds of lung cancer incidence.

Strata		Number of Subjects cases/controls	MPG activity and odds of lung cancer incidence			
			OR^a (95% CI)	p value	$OR_{adj}{}^a$ (95% CI)	p value
All Subjects		51 / 88	**1.40** **(1.10 – 1.78)**	**0.007**	**1.41[b]** **(1.10 – 1.81)**	**0.007**
Smoking Status	**Never**	3 / 8	2.15 (0.44 – 10.58)	0.35	-	-
	Current	19 / 30	**1.58** **(1.03 – 2.42)**	**0.04**	**1.61[b]** **(1.03 -2.53)**	**0.04**
	Former	29 / 50	1.27 (0.93 – 1.72)	0.13	1.29[b] (0.94 – 1.77)	0.11
Gender	**Male**	35 / 47	1.12 (0.83 – 1.52)	0.46	1.10[c] (0.81-1.50)	0.55
	Female	16 / 41	**2.08** **(1.30-3.35)**	**0.002**	**1.89[c]** **(1.15-3.08)**	**0.01**

[a]odds of lung cancer incidence [b]adjusted for age, sex and smoking duration
[c]adjusted for age and smoking duration

Stratification for smoking status revealed an increased adjusted odds ratio for lung cancer in current smokers that was significant (OR_{adj} 1.61 95% CI 1.03-2.53, p = 0.04) and an increased adjusted odds ratio in former smokers that was not significant (OR_{adj} 1.29 95% CI 0.94-1.77, p = 0.11). Stratification for gender revealed a significant increased adjusted

odds ratio of lung cancer in female subjects (OR$_{adj}$ 1.89, 95% CI 1.15-3.08, p = 0.01), but not in male subjects (OR$_{adj}$ 1.10, 95% CI 0.81-1.50, p = 0.55).

A second binary logistic regression analysis was performed with MPG activity entered as a categorical variable. To illustrate the size of effect of MPG activity on lung cancer incidence in females, study subjects were divided according to median activity of controls (Table 6.22). The adjusted odds ratio of lung cancer incidence was 6.33 (95% CI 1.17–34.15, p = 0.03) for female subjects above the median MPG activity compared to females with below median activity.

Table 6.22 The odds of lung cancer in female cases and controls, stratified according to median MPG activity of controls.

MPG Repair Activity in females fmole/µg DNA/hr	Cases n (%)	Controls			
		n (%)	Unadjusted[a]		Adjusted[ab]	
			OR (95% CI)	p value	OR (95% CI)	p value
< 3.5	2 (12.5)	21 (51.2)	1	-	1	-
≥ 3.5	14 (87.5)	20 (48.8)	7.35 (1.48–36.52)	0.02	6.33 (1.17–34.15)	0.03

[a]odds of lung cancer incidence (reference = below median MPG activity)
[b]OR adjusted for age and smoking duration

222

6.1.4 A Comparison of MGMT and MPG activity in Nuclear PBMC Samples

The results for both DNA repair assays were available in 119 subjects, this included 77 controls (64.7%) and 42 cases (35.3%). The mean age of the combined DNA repair protein study population was lower than the mean age of the whole study population (p = 0.06), there were no other significant differences (Table 6.23). The repair activities of MGMT and MPG were found to be significantly correlated in nuclear PBMC samples (p = 0.0001, R^2 0.26) (Figure 6.15).

Table 6.23 Comparison of combined and whole study population

	Whole Study n = 471	Combined Study n = 119	p value
Cases / Controls (% cases)	164 / 307 (34.8)	42 / 77 (35.3)	0.90
Age mean (yrs) ± sd	67.0 ± 10.5	65.4 ± 11.4	0.06
Sex M / F (% male)	297 / 174 (63.1)	68 / 51 (57.1)	0.12
Alcohol mean (units/wk) ± sd	15.6 ± 34.9	13.0 ± 31.8	0.68
Ever Smoked n = yes (% yes)	441(93.6)	110 (92.4)	0.54
Current Smokers n = yes (%[b])	168 (38.1)	42 (35.3)	0.98
Smoking Duration[a] mean (yrs) ± sd	42.1 ± 13.4	41.2 ± 14.0	0.39
Packyears[a] mean ± sd	52.0 ± 34.0	51.2 ± 31.5	0.78
Age Started Smoking[a] mean (yrs) ± sd	16.5 ± 4.3	16.2 ± 3.2	0.33
Cigarettes / day[a] mean ± sd	24.0 ± 13.8	24.1 ± 12.6	0.94

[a]in ever smokers only

$y = 2.4606x + 3.4543$

$R^2 = 0.2587$
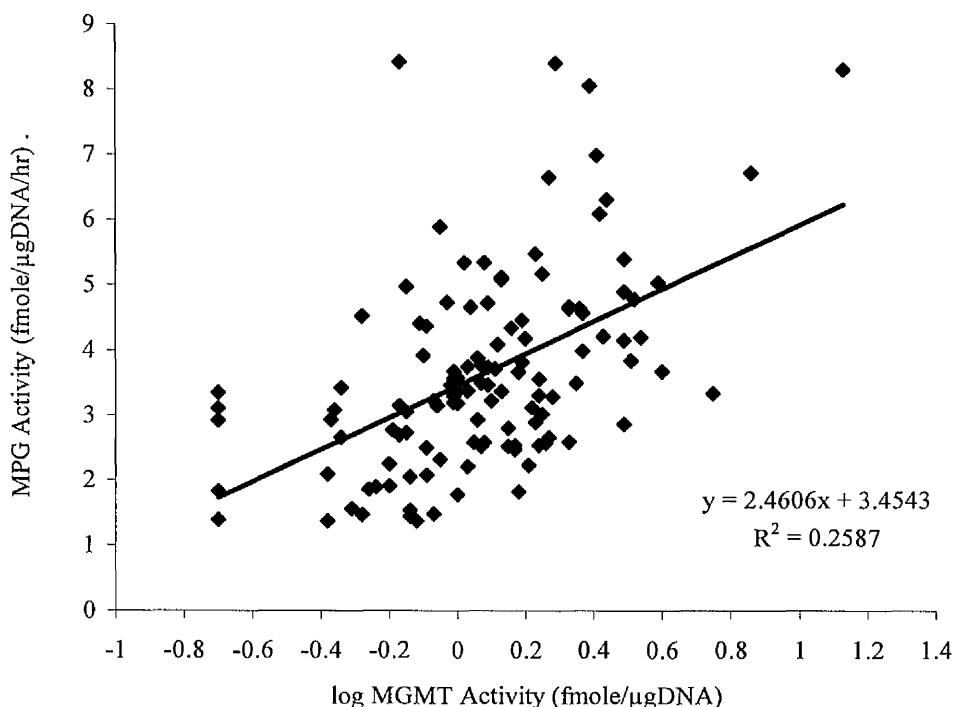
Figure 6.15 Correlation of MPG and log MGMT activity.

6.1.4.1 MGMT and MPG Activity and Case Status

The relationship between MPG and MGMT activity was investigated after stratifying subjects according to case status. The significant positive correlation seen in the whole population was seen in both cases and controls (Figure 6.16).

(a) Cases
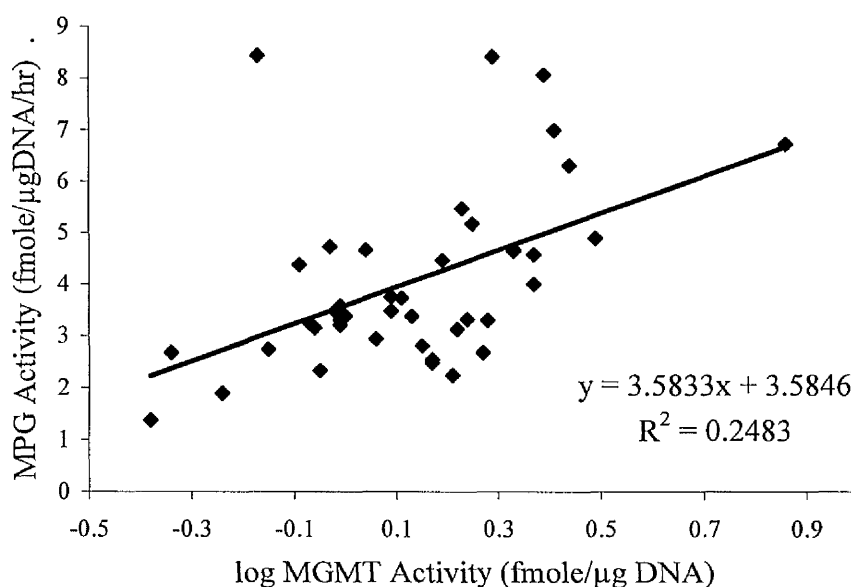
(b) Controls
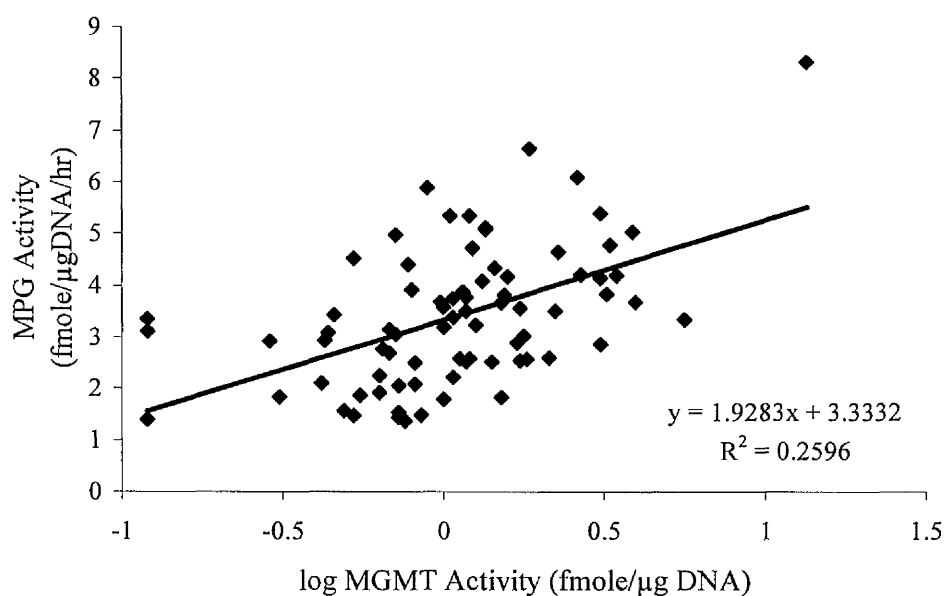
Figure 6.16 MPG and log MGMT activity in: (a) cases and (b) controls.

6.1.4.2 MGMT and MPG Activity and Age

The correlation in MGMT and MPG was examined after the study population was divided according to the median age of controls. The correlation for both categories was significant, however the relationship was stronger in the older age group (R^2 0.30, p = 0.0005 vs. R^2 0.18, p = 0.001) (Figure 6.17).

(a) less than median age of controls (66 or less)

$$y = 1.9076x + 3.4485$$
$$R^2 = 0.176$$
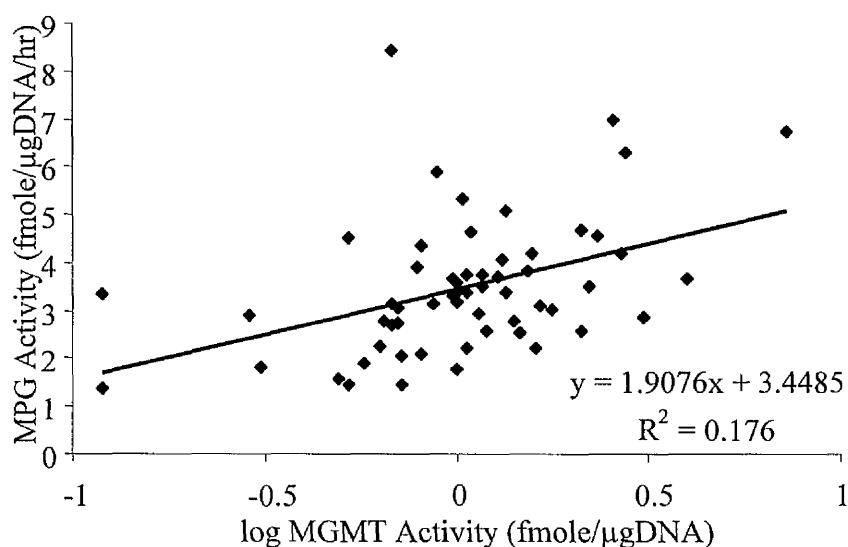
(b) greater than median age of controls (above 66)

$$y = 2.6923x + 3.453$$
$$R^2 = 0.3014$$
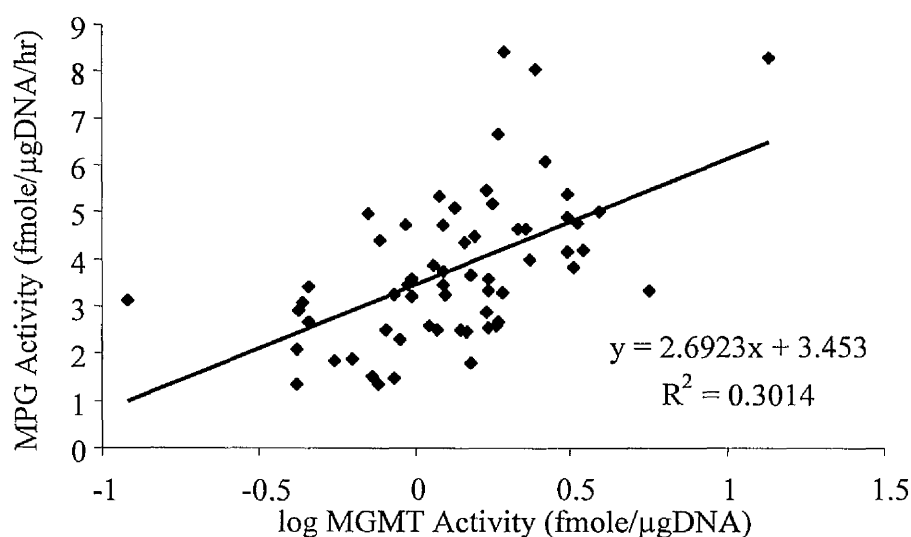
Figure 6.17 MPG and MGMT activity in subjects stratified according to age: (a) below and (b) above the median age of controls.

6.1.4.3 MGMT and MPG Activity and Gender

The correlation of MGMT and MPG was examined after the study population was divided according to gender. The correlation for both males and females was significant and similar to the correlation seen for the study as a whole (Figure 6.18).

(a) Males

$y = 2.4401x + 3.478$

$R^2 = 0.2548$

(b) Females

$y = 2.2208x + 3.4519$

$R^2 = 0.2377$
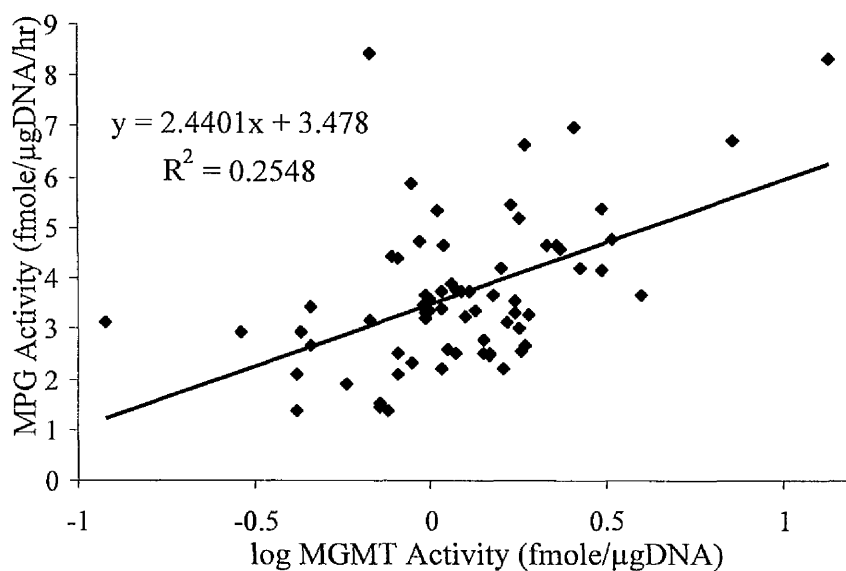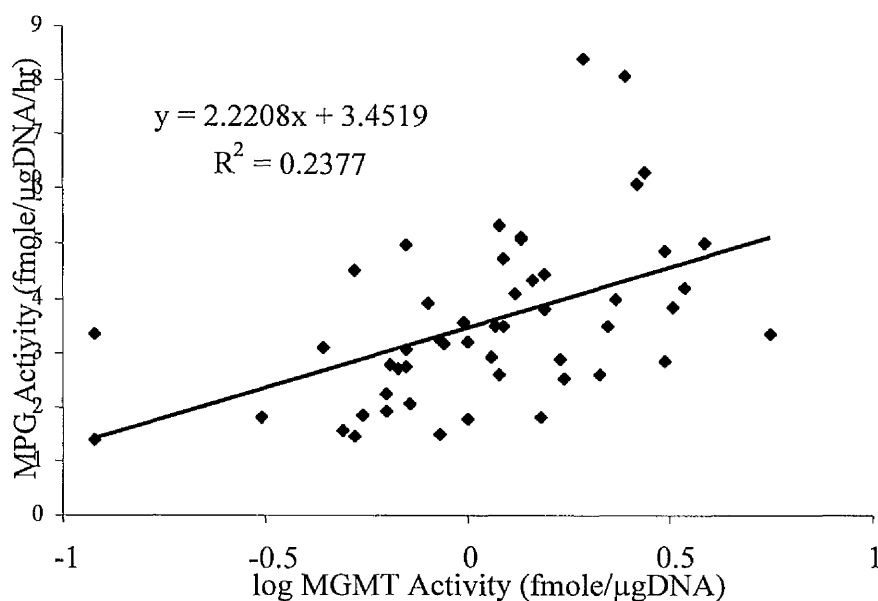
Figure 6.18 Correlation of MPG and log MGMT activity in: (a) males and (b) females.

6.1.4.4 MGMT and MPG Activity and Smoking Status

The effect of smoking status was analysed with respect to MPG and MGMT activity levels. MPG and MGMT activity levels were not correlated in never smokers (R^2 0.02, p = 0.82), there was a weak correlation in current smokers (R^2 0.09, p = 0.04) and a much stronger correlation in former smokers (R^2 0.42, p = 0.0002) (Figure 6.19).

(a) Former smokers

$$y = 2.8729x + 3.3956$$
$$R^2 = 0.4181$$

(b) Current smokers

$$y = 1.6204x + 3.694$$
$$R^2 = 0.0903$$
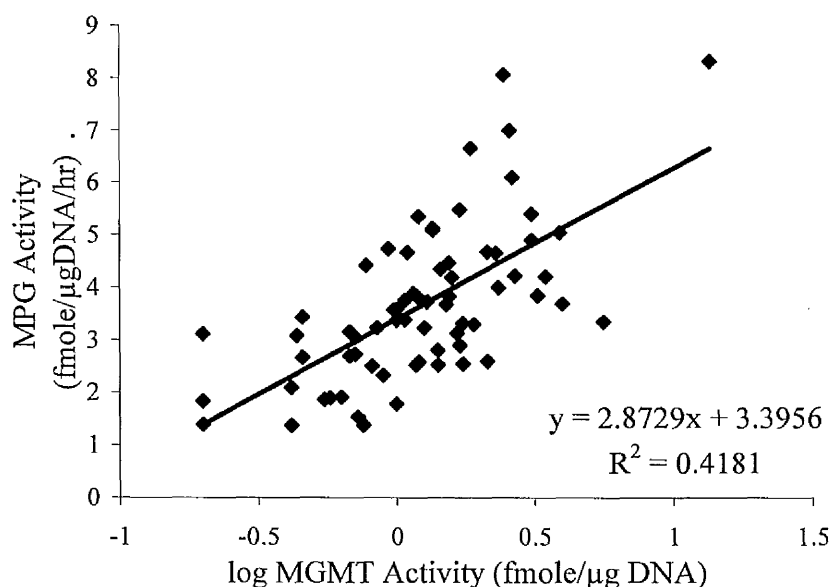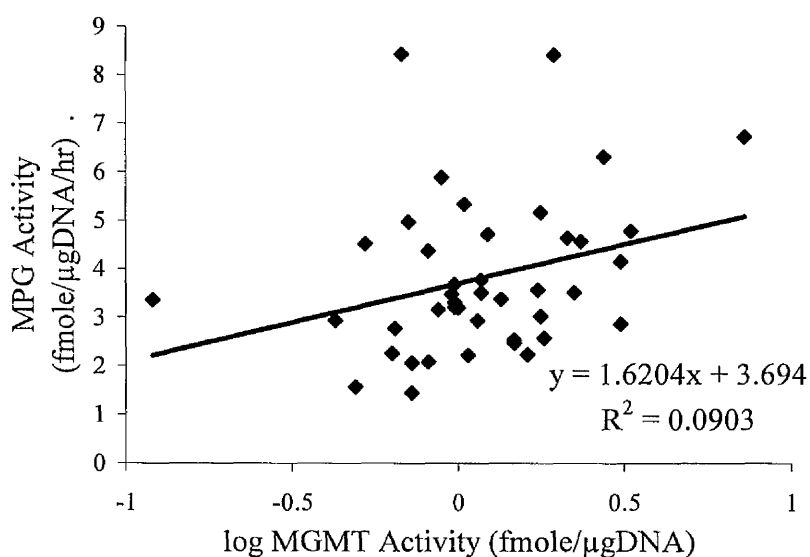
Figure 6.19 Correlation of MPG and log MGMT activity in: (a) former and (b) current smokers.

6.1.4.5 MGMT and MPG Activity and DNA Extraction

The mean amount of DNA extracted from the 243 PBMC samples assayed for MPG and MGMT activity was $65.7 \pm 35.5\mu g$ (range $8.9 - 348$). The yield was significantly greater in cases than controls (76.4 ± 47.7 vs. 60.2 ± 25.7 μg DNA, $p = 0.005$), there was no association with age (R^2 0.01, $p = 0.09$) or gender. Current smokers had a higher yield than former or never smokers but this did not reach significance ($p = 0.10$) (Table 6.24).

Table 6.24 PBMC DNA quantification according to gender, case and smoking status.

Strata		Number Of Subjects	Mean DNA Quant. μg DNA $\pm$ sd.	p value
Gender	Male	147	63.2 ± 31.2	0.17
	Female	96	69.6 ± 41.2	
Case Status	Case	83	76.4 ± 47.7	**0.005**
	Control	160	60.2 ± 25.7	
Smoking Status	Never	18	59.1 ± 30.3	0.10
	Former	145	62.8 ± 35.6	
	Current	80	72.7 ± 35.9	

A linear regression analysis of mean sample DNA extraction and repair activity of both MPG and MGMT in samples with accepted values was performed. There was no association between amount of DNA extracted and log MGMT activity (R^2 0.001, $p = 0.60$). The association between amount of DNA extracted and MPG activity did approach significance, but was not judged to be of practical significance because of the low R^2 value (R^2 0.02, $p = 0.07$) (Figure 6.20).
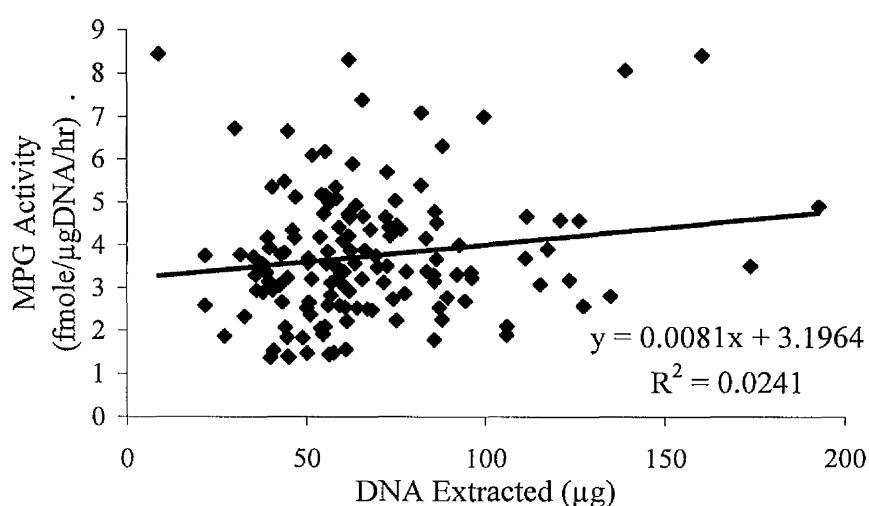
Figure 6.20 PBMC sample DNA extraction and MPG repair activity.

6.1.5 Discussion

The aim of this study was to investigate whether differences in DNA repair activity for two repair mechanisms of alkylation damage, MGMT and MPG, existed between lung cancer cases and controls.

The results of MPG analysis showed MPG repair activity to be significantly greater in lung cancer cases than controls (p = 0.005). This is the first report of an association between MPG repair activity and lung cancer and indeed the first report to find an association between MPG repair activity and any form of cancer in humans. The finding of increased MPG repair activity in lung cancer cases is consistent with previous studies that have reported increased susceptibility to alkylating agents in cells overexpressing MPG (reviewed in section 1.4) (Coquerelle, Dosch et al. 1995; Calleja, Jansen et al. 1999; Fishel, Seo et al. 2003; Rinne, Caldwell et al. 2004). Increased MPG mRNA expression has also been reported in breast cancer cells lines (Cerda, Turk et al. 1998) and tumour tissue compared to normal controls (Kim, Lee et al. 1998; Kim, An et al. 2002; Kim, Ahn et al. 2003). The association of increased DNA repair activity with an increased risk of lung cancer is in direct contrast to similar studies of the bifunctional DNA glycosylase hOGG1, which have shown reduced repair activity in PBMCs to be a risk factor for lung cancer (Gackowski, Speina et al. 2003; Paz-Elizur, Krupsky et al. 2003). This may be due in part to the unusually broad substrate range of the MPG enzyme compared to other DNA glycosylases (Singer and Hang 1997). It has been postulated that high levels of MPG activity might increase susceptibility to cancer by creating an imbalance in the base excision repair pathway. Relatively benign products of alkylation exposure such as N7-meG are rapidly repaired by MPG resulting in the creation of toxic intermediaries of the BER pathway (Rinne, He et al. 2005). The presence of increased quantities of AP sites, which are cytotoxic and mutagenic (Dianov, Sleeth et al. 2003; Yu, Lee et al. 2003) could therefore create an environment where the probability of mutations is higher and as a consequence cancer is more likely to develop.

One interesting observation was that lung cancer risk was increased in females with increasing MPG activity but not in males. There was a six fold increased risk of lung cancer in females with above median MPG repair activity compared to females with below median activity (p = 0.03). The reason for such a marked and gender specific effect is not clear. It remains controversial whether there is a gender difference in tobacco smoke

carcinogen susceptibility with epidemiological studies reporting inconsistent results; some reporting increased susceptibility in females (Risch, Howe et al. 1993; Zang and Wynder 1996; Henschke and Miettinen 2004) and others reporting no difference (Kreuzer, Boffetta et al. 2000; Bain, Feskanich et al. 2004). Comparison of male and female cases from this study revealed that male cases had a significantly higher packyear exposure than female cases (59.1 ± 30.2 vs. 41.4 ± 16.7 packyears, p = 0.0007), but the duration of smoking was not significantly different between the groups (46.2 ± 12.2 vs. 46.5 ± 11.1 years, p = 0.88). A separate but possibly more important issue is whether there are gender differences in the biology of lung cancer development. The fact that differences do exist is most clearly demonstrated in the proportion of histological types of lung cancer seen in males and females; squamous cell carcinoma is the most common cell type in males and adenocarcinoma in females (Janssen-Heijnen and Coebergh 2003). In this study, females were twice as likely to be diagnosed with adenocarcinoma and half as likely to be diagnosed with squamous cell carcinoma as males.

At a molecular level, differences between males and females have been reported with respect to lung cancer. For example, smoking induced CYP1A1 mRNA levels have been reported to be higher in lung tissue from females than males (Mollerup, Berge et al. 2006) and in a connected study female lung tissue had higher levels of bulky DNA adducts (Mollerup, Ryberg et al. 1999). Polymorphisms in CYP1A1 and GSTM1 have also been reported to increase lung cancer risk in females over males (Dresler, Fratelli et al. 2000) and females are also more likely to carry a mutation in the K-ras oncogene (Nelson, Christiani et al. 1999; Ahrendt, Decker et al. 2001). Differences in NNK susceptibility have also been reported (Hill, Affatato et al. 2005). There is also evidence for differences in outcomes after the use of certain chemotherapeutic agents in the treatment of lung cancer (Thomas, Doyle et al. 2005). One example is seen with the epidermal growth factor receptor (EGFR) inhibitor gefitinib, which is more effective in females than males (Janne, Gurubhagavatula et al. 2004), possibly due to the increased frequency of EGFR mutations in females and in patients with adenocarcinoma (Paez, Janne et al. 2004). PARP inhibition may also have gender specific effects with increased benefit in male compared with female mouse models (Szabo, Pacher et al. 2006). This mechanism might be due to the oestrogen attenuation of poly(ADP-ribosyl)ation or due to complexes formed with oestrogen receptor α (Szabo, Pacher et al. 2006). There is therefore some evidence for biological variance in lung cancer development between the sexes. However, it is not clear why a difference should exist for MPG repair activity.

Hormonal differences between men and women might be a factor. Oestrogen receptors (ER) are expressed in healthy lung and lung tumours (Dubey, Siegfried et al. 2006). Exposure to oestrogen results in ERβ mediated cellular proliferation of lung cancer cells (Omoto, Kobayashi et al. 2001; Stabile, Davis et al. 2002). Studies examining the effect of exogenous and endogenous oestrogens on lung cancer incidence have produced conflicting results. Taioli *et al*, reported a reduction in risk of adenocarcinoma with early menopause but increased risk with the use of HRT (Taioli and Wynder 1994). Two further studies have however shown a protective effect of HRT on the risk of lung cancer development (Kreuzer, Gerken et al. 2003; Schabath, Wu et al. 2004). A recent retrospective study by Ganti *et al*, has shown that postmenopausal women receiving HRT were diagnosed with lung cancer at a significantly earlier age and had a significantly worse outcome than postmenopausal women not receiving HRT (Ganti, Sahmoun et al. 2006). Conflicting results may be due to the confounding effects of progesterone in HRT which is thought to inhibit lung cancer growth and the antioestrogenic effects of smoking (Dubey, Siegfried et al. 2006). A potential clinical effect of oestrogen inhibition was seen in a population study of breast cancer survivors. One treatment group given the aromatase inhibitor exemestane, showed a 67% reduction in the incidence of primary lung cancer compared with a tamoxifen treated group (Coombes, Hall et al. 2004). Tamoxifen shows partial ER agonist properties in certain tissues.

MPG itself has been shown to interact directly with the nuclear transcription factor, oestrogen receptor alpha (ERα). A study by Likhite *et al*, showed that the presence of ERα was able to increase MPG binding to hypoxanthine adducts in a double stranded oligonucleotide up to 2.5 fold and that hypoxanthine cleavage was, as a consequence, significantly enhanced (Likhite, Cass et al. 2004). It is not clear if this mechanism could explain increased MPG activity in the lung as ERβ is thought to be the dominant receptor in the lung (Hershberger, Vasquez et al. 2005).

Caution must be used when interpreting the results of case-control studies because of the difficulty in defining the temporal relationship between significant findings and the outcome being investigated. Clearly the presence of lung cancer itself may have caused an increase in MPG repair activity. However, it is noteworthy that MPG activity was not increased in male cases compared to controls and repair activity was not related to disease stage in NSCLC (Table 6.15). Further studies are needed to validate these findings; to determine causality a prospective study would be required. The possibility that increased

MPG activity may play a causal role in lung cancer development must still be considered. This study did not measure MPG activity in the target cells of lung cancer development, namely bronchial epithelial cells. Attempts were made to measure MPG in lung tissue and correlate findings with PBMCs but the non-specific cleavage seen in lung was prohibitive to further investigation. MPG activity has never been measured in BECs and PBMCs from the same individual. It is therefore unknown if there is correlation between different tissues. A strong correlation between PBMCs and lung tissue has been shown in repair activity of another DNA-glycosylase hOGG1 (Paz-Elizur, Krupsky et al. 2003). MPG activity has only once previously been examined in human samples prior to this study (Hall, Bresil et al. 1993). Hall *et al*, showed that MPG activity was higher in current smokers (n = 20) compared to non-smokers (n = 17). A bimodal distribution was seen in repair activity that was independent to smoking status (Hall, Bresil et al. 1993). The distribution of MPG repair activity in this study was unimodal; the previously seen bimodal distribution may have simply been a consequence of smaller numbers. The finding that MPG activity was higher in smokers is consistent with findings that a methylating agent (DMN) induced MPG activity in the livers of rodents (Hall, Bresil et al. 1990). Activity in ever smokers was higher than never smokers, a difference that approached significance (p = 0.07). However, there was no significant difference between current and former smokers. This may be a consequence of misclassification of smoking status due to inaccurate self reporting, or due to an upregulation of the MPG gene in ever smokers to a level that does not return to non-smoking levels. This study was not designed to directly investigate the effect of smoking on DNA repair activity; this would have to be addressed in a future study using healthy smokers and non-smokers, to avoid the confounding effect of both benign and malignant respiratory disease. Cases and controls were re-analysed according to smoking status. Although the activity of MPG was greater in cases in all smoking categories, only current smoking cases had significantly higher MPG activity than controls (p = 0.05).

Due to an intermittent problem with the MPG assay, only 60% of samples assayed resulted in acceptable results of MPG activity. This was far from ideal, however, cases and controls were lost in equal proportion and the rejection of samples was essentially random. Comparison of the MPG study population with the whole study population revealed no significant differences in demographic details or smoking exposure and was therefore judged to be representative of the study as a whole.

The MGMT assay was more reliable and almost all samples analysed resulted in acceptable data. Nuclear MGMT activity has never previously been measured with respect to lung cancer susceptibility in PBMCs. The primary conclusion from this aspect of the study was that MGMT nuclear activity in the whole study population was not related to lung cancer case status. This finding is consistent with two previous case control studies that examined MGMT activity with respect to lung cancer in whole PBMC extracts (Boffetta, Nyberg et al. 2002; Margison, Heighway et al. 2005). An increased risk of tumour induction with reduced levels of MGMT is biologically plausible and was the prestudy hypothesis. Evidence from murine models indicate that MGMT null mice exposed to alkylating agents are more susceptible to tumour induction (Tsuzuki, Sakumi et al. 1996; Iwakuma, Sakumi et al. 1997; Sakumi, Shiraishi et al. 1997) and mice over-expressing MGMT are less susceptible to alkylating agents including NNK (Liu, Qin et al. 1999; Zhou, Manguino et al. 2001).

The lack of association with lung cancer seen in this study and previous studies of PBMCs might be due to several factors. PBMCs are not the target cell in lung cancer development and are not directly exposed to carcinogens deposited on to the lung's epithelial lining. Activity measured in PBMCs does not correlate with activity in bronchial epithelial cells (O'Donnell, Barber et al. 1999) and therefore may not relate to the cells directly involved in lung cancer development. However, MGMT activity measured in bronchial epithelial cells has also been shown not to be associated with lung cancer status (Povey, O'Donnell et al. 2006). Technically, the harvesting of bronchial epithelial cells is difficult and invasive, requiring patients to undergo bronchoscopy. The use of PBMCs allows samples to be taken on larger numbers of individuals and is much less invasive for patients. The use of surrogate tissue is appropriate in the setting of detecting potential biomarkers for lung cancer susceptibility; as samples would need to be taken from individuals who are well and therefore not due to undergo invasive medical investigations.

The large inter-individual variation seen in MGMT repair activity (46.1 fold) seen in this study is consistent with previous reports (Margison, Povey et al. 2003) and may be one reason why differences between cases and controls are difficult to detect. A 3.5 fold variation in activity has been described in the same individuals over a several week period (Janssen, Eichhorn-Grombacher et al. 2001) and a smaller variation over a 24 hour period (Marchenay, Cellarier et al. 2001). These studies have examined whole cellular extracts; it is not known whether the same variation is found in the nuclear protein. The wide intra

and inter-individual variation in MGMT activity makes the use of genotyping as a more stable guide to MGMT status appealing. Previous work by Margison *et al*, showed repair activity to be associated with polymorphisms in codon 178 and intron 1 of the MGMT gene (Margison, Heighway et al. 2005). The majority of the MGMT study population were successfully genotyped for both codon 178 (99.6%) and intron 1 (98.2%) polymorphisms. No significant association was found between genotype and MGMT repair activity both in the whole population and after categorisation for case status. The reason for this discrepancy in results might be related to the types of samples used. In this study a nuclear fraction of PBMCs was used and in the previous study whole PBMC extracts were assayed. The level of MGMT available for repair in the nucleus at any one time may rely on numerous factors that do not influence the level of MGMT produced by the cell as a whole. Genotype may therefore be more closely related to whole cell extract than the small proportion in the nucleus.

A lack of association was seen between multiple measures of self reported smoking exposure and MGMT activity in the study population as a whole. This is similar to previous studies both of PBMCs and other tissues (reviewed in section 1.4.1). The population was stratified according to smoking status and MGMT activity re-examined between cases and controls. Current smoking cases had a significantly greater activity than both current smoking controls ($p = 0.01$) and former smoking cases ($p = 0.05$). The accuracy of smoking status between former or never smokers could be questioned and might have resulted in a degree of misclassification of smoking status, but it is highly unlikely that subjects would falsely claim to be current smokers if they were not. As previously discussed, the observation of a significant finding in a case-control study does not indicate a causal relationship but simply an association. The relatively high level of MGMT repair activity in current smoking cases may reflect an upregulation of MGMT caused by the presence of a serious disease and ongoing exposure to tobacco smoke.

When MGMT and MPG values were analysed in the 119 subjects who had acceptable results for both, it was found that there was a significant positive association between the two repair proteins ($R^2 = 0.26$, $p = 0.0001$). The only previous study to have examined MPG and MGMT activity found no significant association (Hall, Bresil et al. 1993). This was a much smaller study analysing whole PBMC extracts and may have lacked the power to detect an association. The finding of a correlation using a nuclear fraction in this study

is supported by the same finding using whole PBMC cellular extract reported in Chapter 5.

There is only limited knowledge on the regulation and expression of MPG. The transcription factor AP-2 is also thought to influence MPG expression (Cerda, Chu et al. 1999). A study by Bouzianne *et al*, showed that MPG expression was increased threefold in the G_1 phase of the cell cycle and fell again after mitosis (Bouziane, Miao et al. 2000). In contrast the expression of MGMT was not affected by the cell cycle when measured in the same study. MGMT activity and expression levels have been shown to be influenced by several factors including genetic polymorphisms, transcription factors e.g. AP-1 and glucocorticoids, promoter methylation, chromatin structure and phosphorylation (Margison, Povey et al. 2003). Evidence from experiments in rat liver have suggested that MGMT and MPG are not co-regulated (Grombacher and Kaina 1995). This study measured repair activity of both mechanisms in a nuclear fraction of PBMCs. An important factor in the correlation seen may not simply be the level of either protein produced in the whole cell, but the mechanism of protein translocation and retention in the nucleus, a two step process seen with MGMT (Lim and Li 1996). Further work is required to investigate these findings.

Chapter Seven

MGMT Allelic Expression Imbalance

7.1   MGMT Allelic Expression Imbalance in Whole Blood

The aim of this study was to investigate the presence and extent of allelic expression imbalance for the MGMT gene and to investigate factors that may influence the imbalance. A polymorphism in codon 178, involving a lysine (AAG) (K allele) to arginine (AGG) (R allele) substitution, was chosen as a marker for allelic expression imbalance.

7.1.1   Study Population

An aliquot of whole blood was collected and stored in RNAeasy consecutively from 60 chest clinic controls and 21 surgical patients. These subjects were genotyped to identify heterozygotes for codon 178 of the MGMT gene. There were 12 heterozygotes identified from chest clinic subjects (20.0%, n = 12 / 60) and 10 from surgical subjects (47.6%, n = 10 / 21). All chest clinic controls were cancer free and all surgical patients were diagnosed with primary lung cancer. A comparison of ages, gender, smoking and recent alcohol exposure between the two groups is detailed below (Table 7.1). Controls reported smoking significantly more cigarettes than cases (p = 0.009), the duration of smoking was on average 5 years longer in the cases, but calculated packyears was significantly higher in controls (p = 0.03).

Table 7.1 Age, gender, alcohol intake and smoking exposure of cases and controls.

Variable	Cases (n = 10)	Controls (n = 12)	p value
Gender m / f (%m)	6 / 4 (60.0)	5 / 7 (41.7)	0.39
Mean age (yrs ±sd)	67.4 ± 6.2	67.5 ± 9.4	0.98
Smoking hx: Current / former / never	5 / 5 / 0	6 / 5 / 1	0.63
Smoking duration[a] (yrs ±sd)	45.1 ± 9.9	40.4 ± 15.6	0.43
Age started smoking[a] (yrs ±sd)	16.5 ± 4.8	16.8 ± 3.4	0.86
Cigarettes / day[a] (±sd)	15.0 ± 5.9	34.6 ± 19.9	**0.009**
Packyears[a] (±sd)	33.8 ± 16.3	70.6 ± 45.7	**0.03**
Alcohol (units/week ±sd)	13.5 ± 13.2	9.9 ± 24.0	0.68

[a] in ever smokers only

7.1.2 Results

7.1.2.1 Definition of Allelic Expression Imbalance

The ratio of alleles in genomic DNA after PCR amplification should in theory be 1:1, as two copies of the gene are present. However PCR amplification and digestion may not necessarily be equally efficient for both alleles. The allelic ratio of genomic DNA, after PCR amplification and digestion, was therefore calculated for all samples. The mean ratio $\pm$ 3 standard deviations was taken as a normal range; values outside this range would therefore be considered to be imbalanced when c-DNA was examined. The mean ratio was 1.1 ± 0.13. The range of acceptable values as defined by the mean $\pm$ 3 standard deviations was 0.71 to 1.49 (Table 7.2).

7.1.2.2 Relative Allelic Expression in Whole Blood

Whole blood samples from 22 heterozygous subjects were processed and the relative allelic expression of the MGMT codon 178 alleles measured on at least one occasion in 21 subjects (95.5%, n = 21 / 22). The relative allelic expression of the single nucleotide polymorphism of codon 178 was measured on three separate occasions. The first assay (1) was independent of the later two assays from mRNA separation onwards; the latter two assays (2+3) used the same c-DNA. The CV of all the repeat assays was less than 15% for each sample.

Allelic expression imbalance was detected in 14 subjects (66.7%, n = 14 / 21). The K allele was most commonly over expressed, representing 12 / 14 of the subjects with imbalance (85.7%) and 57.1% of subjects overall. The R allele was over-expressed in only two subjects, this represented 2 / 14 of the subjects with imbalance (14.3%) and 9.5% of subjects overall. The distribution of allelic expression imbalance was not significantly different according to gender, case or smoking status (Table 7.3). The degree of imbalance was up to a maximum of 3.7 fold for over-expression of the K allele and 2.5 fold for over-expression of the R allele.

Table 7.2 Genomic and c-DNA allelic ratios for all study subjects.

Study Subject	Genomic DNA Ratio R:K	c-DNA Ratio R:K			Mean c-DNA Ratio R:K	
		1	2	3	mean	± sd
S8	1.2	1.3	1.2	1.2	1.2	0.06
S9	1	2.2	1.9	1.9	**2.0**	0.15
S10	1	1.8	1.5	1.4	**1.6**	0.21
S11	1.1	1.3	1.3	1.3	1.3	0
S19	1.2	2.1	a	a	**2.1**	-
S23	1	3.4	3.3	3.4	**3.4**	0.05
S25	1.2	1.9	1.9	1.8	**1.9**	0.06
S26	1.1	1.2	1.2	b	1.2	0.00
S28	1.1	1.1	1.3	b	1.2	0.12
S29	1.1	0.8	0.9	0.9	0.9	0.06
442	1	-	3.7	3.6	**3.7**	0.07
444	1	1.8	1.9	1.7	**1.8**	0.10
459	0.7	1.2	1.2	1.2	1.2	0
462	1.2	1.5	1.8	1.7	**1.7**	0.15
470	1.3	2.2	2.6	2.5	**2.4**	0.21
471	1.3	2	2	2.1	**2.0**	0.06
487	1.2	0.6	0.6	0.7	**0.6**	0.05
489	1.1	1.4	1.6	1.5	**1.5**	0.10
491	1.1	c	c	c	-	-
492	1.1	1.6	1.5	1.6	**1.6**	0.06
493	1.1	0.9	0.9	0.9	0.9	0
498	1.1	0.4	0.5	0.4	**0.4**	0.06
Mean ± 3 sd	1.1 ± 0.39 (range = 0.71 – 1.49)					

[a]insufficient RNA, [b]insufficient c-DNA, [c]RNA extraction failed

Table 7.3 Comparison of allelic expression imbalance in all subjects and according to case status, gender and smoking status.

Category (n)	Allelic Expression (%)			p value
	K > R	**K = R**	**R > K**	
All (21)	12 (57.1)	7 (33.3)	2 (9.5)	-
Cases (10)	5 (50.0)	5 (50.0)	0 (0)	0.17
Controls (11)	7 (63.6)	2 (18.2)	2 (18.2)	
Males (11)	5 (45.5)	4 (36.4)	2 (18.2)	0.30
Females (10)	7 (70.0)	3 (30.0)	0 (0)	
Current Smokers (11)	8 (72.7)	2 (18.2)	1 (9.1)	0.28
Non-Current Smokers (10)	4 (40.0)	5 (50.0)	1 (10.0)	
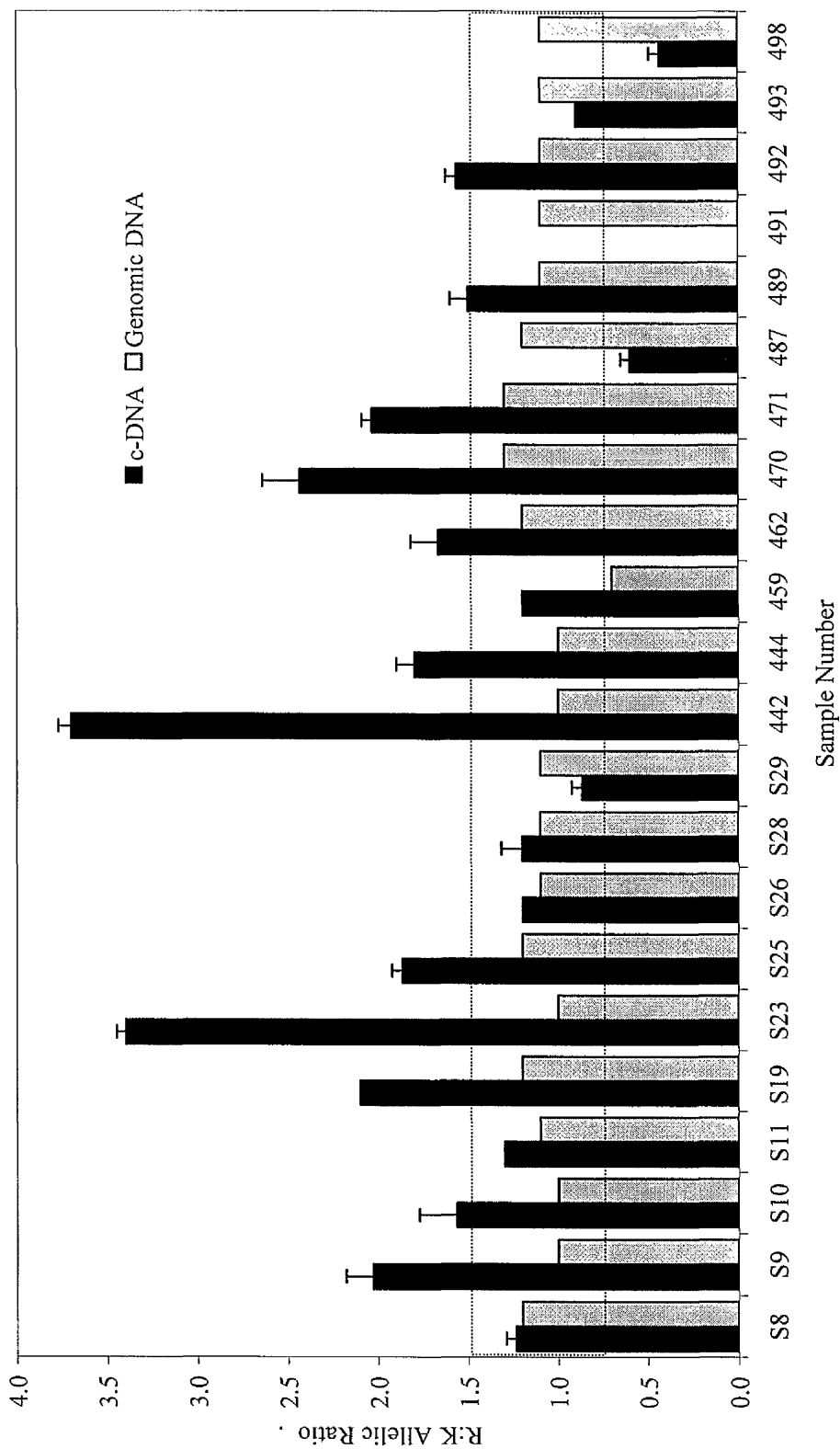
Figure 7.1 Allelic expression imbalance in whole blood (dotted line represents the mean genomic DNA ratio ± 3sd).

240

7.2 MGMT Allelic Expression Imbalance in Whole Blood, Lung and Tumour Tissue

The aim of this study was to investigate the presence and extent of allelic expression imbalance (AEI) in lung and tumour tissue from patients who had undergone surgical resection for treatment of primary lung cancer. The results from lung and tumour tissue were compared to the results from whole blood (Section 7.1).

7.2.1 Study Population

Lung tissue was taken from a macroscopically normal area of resected lung away from the primary tumour in 21 surgical patients and stored in RNAlater RNA stabilisation reagent. Ten subjects were heterozygous for the codon 178 polymorphism of MGMT (47.6%, n = 10 / 21). The demographic details of these 10 subjects have previously been detailed in Table 7.2. Information regarding AEI was available in lung and tumour from 8 out of 10 subjects.

7.2.2 Results

7.2.2.1 Ratio of Genomic Alleles in Whole Blood, Lung and Tumour Tissue

DNA was extracted from all three tissues to be examined and the ratio of codon 178 K and R polymorphisms determined (Table 7.4). The mean ratio in genomic DNA of the K:R allele was 1.25 ± 0.09 in whole blood, 1.25 ± 0.14 in normal lung tissue and 1.13 ± 0.10 in tumour tissue. There was no evidence of the presence of multiple copies of MGMT alleles in tumour tissue. The overall mean value from all tissues was 1.21 ± 0.12; the range of normal was defined as the mean $\pm$ 3 standard deviations and equated to: 0.83 to 1.58.

7.2.2.2 Allelic Expression Imbalance in Whole Blood, Lung and Tumour Tissue

The AEI results for whole blood in this study were in general agreement with the results for whole blood in the previous study (Section 7.1). However, sample S10 in this study did not show AEI (ratio 1.3); in the previous study borderline imbalance was recorded (ratio 1.6). AEI was found in 4 out of 8 whole blood samples, all overexpressing the K allele.

Table 7.4 Genomic allelic ratios for all study subjects in whole blood, normal lung and primary lung tumour tissue.

Study Subjects	Genomic DNA R:K ratio		
	Whole Blood	Normal Lung	Lung Tumour
S9	1.2	1.4	1
S10	1.4	1.2	1.1
S11	1.2	1.2	1.3
S19	1.2	1.3	1.2
S23	1.2	1.1	1.1
S25	1.2	1.5	1.2
S26	1.4	1.2	1
S28	1.2	1.1	1.1
Mean ± sd	1.25 ± 0.09	1.25 ± 0.14	1.13 ± 0.10
Range (± 3sd)	0.97 – 1.53	0.83 – 1.67	0.81 – 1.44

Table 7.5 c-DNA allelic ratios for all study subjects in whole blood, normal lung and primary lung tumour tissue.

Study Subjects	c-DNA R:K ratio		
	Whole Blood	Normal Lung	Lung Tumour
S9	**2.1**	**2.2**	**3.5**
S10	1.3	1.2	0.8
S11	1.2	1.3	1.4
S19	**2.1**	**2.7**	**2.8**
S23	**3.1**	**3.1**	**8.2**
S25	**1.9**	**2.0**	**3.1**
S26	1.2	1.2	1.2
S28	1.2	1.5	1.5

There was no imbalance detected in the other 4 samples. Detection of AEI was a consistent finding whichever tissue was examined (Table 7.5) (Figure 7.2). The direction and degree of imbalance was also a consistent finding between whole blood and lung tissue. Tumour tissue showed imbalance in the same direction as blood and normal lung but the degree of imbalance was greater in three out of four samples. In sample S23 for example, the allelic ratio was 3.1 in both whole blood and normal lung but 8.2 in tumour tissue. The genomic DNA allelic ratio in tumour tissue was 1.1 in this sample.
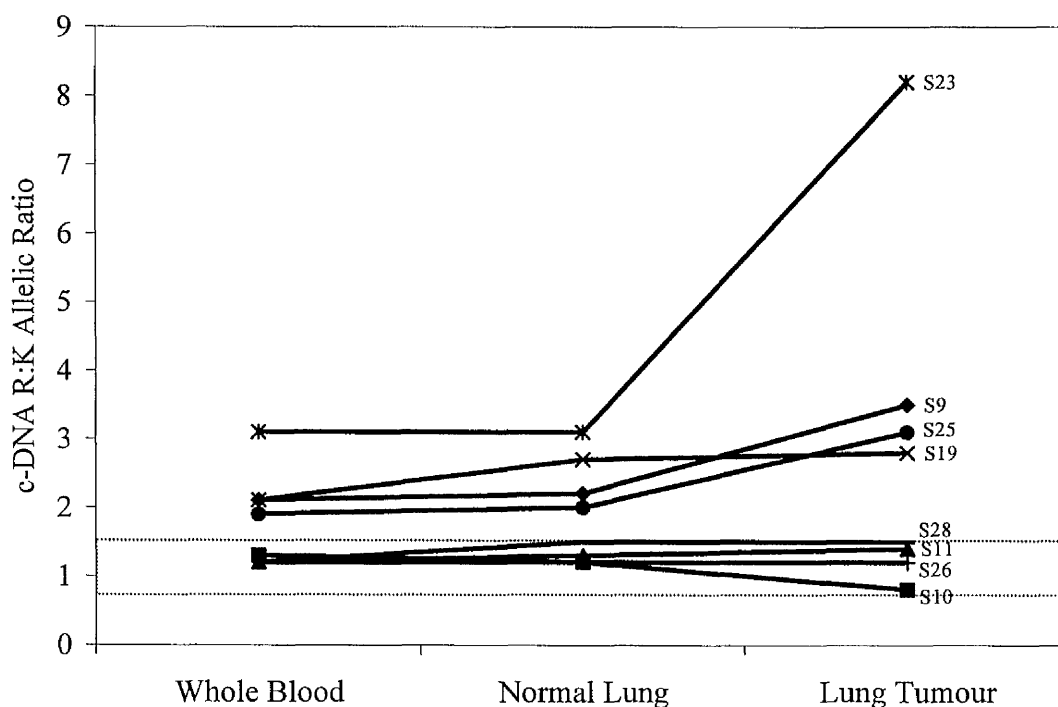
Figure 7.2 Comparison of AEI between whole blood, normal lung and lung tumour tissue (dotted line represents mean R:K ratio in genomic DNA ± 3 sd).

7.3 Discussion

Allelic expression imbalance (AEI) refers to a measured difference in the ratio between the amounts of genomic DNA and mRNA for each allele of a gene in a certain tissue (Wang and Sadee 2006). Differences in allelic expression have long been observed in oncogenes and tumour suppressor genes in tumour tissue which reflect gross alterations in genomic DNA e.g. amplification, deletion, duplications and translocations (Heighway, Margison et al. 2003). Such changes are fundamental to inducing genetic instability and promoting tumour development (Lengauer, Kinzler et al. 1998). However, allelic expression imbalance is also thought to be a common finding in human genes from normal tissues (Bray, Buckland et al. 2003). In one study of 1389 genes, in the PMBCs of 12 subjects, AEI was detected in at least one individual in 731 or 53% of genes (Pant, Tao et al. 2006). It is hypothesised that differences in allelic expression may contribute to the phenotypic variation seen between individuals and in the case of lung cancer therefore might be a factor in determining susceptibility to lung cancer development.

Factors that effect allelic transcription can be *trans*-acting e.g. transcription factors or *cis*-acting e.g. the binding sites of transcription factors or methylation; cis acting genetic variation can result in unequal modifications to gene expression or mRNA processing and so result in differences of allelic expression, in subjects heterozygous for such changes (Wang and Sadee 2006). *Cis*-acting mechanisms are thought to account for up to a third of interindividual differences in gene expression (Pastinen and Hudson 2004). Imprinting has also been associated with variance of allele expression, but this is thought to apply to 1% of genes (Sleutels, Barlow et al. 2000) and has never been seen in the region of the MGMT gene (Heighway, Margison et al. 2003). Polymorphic variants of the MGMT gene have in some studies been shown to effect the risk of lung cancer development (Kaur, Travaline et al. 2000; Cohet, Borel et al. 2004). The relative expression of different alleles may therefore have effects on phenotype even in individuals with the same genotype, if for instance a 'higher risk' allele is preferentially expressed in a target tissue. MGMT repair activity shows marked variability both between different individuals and between different tissues from the same individual (Margison, Povey et al. 2003). AEI maybe one reason accounting for this variability.

Allelic expression imbalance of the MGMT gene is not a new finding and has previously been reported in normal lung tissue by Heighway *et al* (Heighway, Margison et al. 2003). The authors reported AEI in 7 out of 12 normal lung tissue samples and the extent of imbalance was up to 4 fold. Margison *et al*, have also reported the presence of MGMT AEI in 6 out of 7 PBMC

244

samples (Margison, Heighway et al. 2005). In this study, which also examined the codon 178 polymorphism, four cases overexpressed the K allele and two cases the R allele; the extent of imbalance was up to 3 fold. The findings in the current study are consistent with those previously reported; with 12 out of 21 samples showing AEI and the extent of imbalance being up to 3.7 fold in whole blood.

The measurement of AEI in whole blood and lung simultaneously had not previously been performed for MGMT and it was not known whether changes in one tissue would reflect changes in another tissue. The results from the surgical arm of this study showed that the presence or absence of AEI was a consistent finding between whole blood and normal lung tissue. The degree of detected AEI was also a consistent finding. The measurement of AEI in whole blood reflected accurately the situation in normal lung tissue and would therefore appear to be a reasonable surrogate; although a definitive conclusion would require a larger sample size. A further conclusion that could be drawn from this observation is that whatever factor or factors that work to alter the relative expression of MGMT alleles, work to the same degree in the two different tissues. This suggests that local environmental factors such as the exposure of the lung to tobacco smoke do not affect the relative levels of allelic transcription. However, the absolute levels of allelic expression could be very different between lung and whole blood. This finding would be consistent with a heterozygous polymorphism of a regulatory element of the MGMT gene.

AEI of the MGMT gene was also detected in primary lung tumours. Interestingly the presence and direction of imbalance in the eight tumours analysed was consistent with the results from normal lung and whole blood. The difference in 3 out of 4 tumours where imbalance was detected was an increase in the degree of imbalance seen. This was most evident for sample S23 where imbalance increased from 3.1 in normal lung and blood to 8.2 in tumour. This was not due to changes in allelic copy number as the ratio of genomic DNA approximated to 1:1. The histology of S23 was carcinoid, an uncommon neuroendocrine lung tumour. One previous study has reported that MGMT CpG island methylation occurs in 25% of carcinoid tumours studied (n = 16) (Chan, Kim et al. 2003). Down regulation of one allele by promoter methylation might account for the increase in allelic imbalance seen in that specific tumour.

Chapter Eight

Measurement of Methylation and Oxidation Damage in DNA

8.1 Measurement of Methylation and Oxidative DNA Damage in Peripheral Blood Mononuclear Cells

The primary aim of this study was to determine the relative levels of oxidative and methylation DNA damage in smokers and non-smokers and between lung cancer cases and controls. A secondary aim was to determine whether adduct level was related to DNA repair activity. The measurement of N7-methyldeoxyguanosine (N7-medG) was used as a marker of methylation damage (performed by K Harrison) and 8-oxodeoxyguanosine (8-oxodG) as a marker of oxidative DNA damage (performed by M Carus). The samples used were peripheral blood mononuclear cells (PBMCs) taken from a subset of the case-control study population.

8.1.1 The study population

The study population was made up of 22 consecutively recruited study subjects from the main case-control study. This included 7 cases (2 adenocarcinoma, 2 small cell, 2 squamous cell and 1 NSCLC) and 15 cancer free bronchoscopy controls. There was no significant difference between cases and controls for the measures of age, gender, smoking exposure and recent alcohol intake (Table 8.1).

Table 8.1 Age, gender, alcohol intake and smoking exposure of cases and controls.

Variable	Cases (n = 7)	Controls (n = 15)	p value
Gender m / f (%m)	5 / 2 (73.3)	11 / 4 (71.4)	0.93
Mean age (yrs ±sd)	71.0 ± 8.4	66.9 ± 12.3	0.43
Smoking hx: Current / former / never	4 / 3 / 0	9 / 5 / 1	0.74
Smoking duration[a] (yrs ±sd)	46.4 ± 11.7	43.0 ± 13.1	0.56
Age started smoking[a] (yrs ±sd)	14.4 ± 3.3	17.7 ± 5.9	0.19
Cigarettes / day[a] (±sd)	22.1± 11.5	22.0 ± 11.0	0.98
Packyears[a] (±sd)	53.1 ± 31.4	51.4 ± 38.2	0.92
Alcohol (units/week ±sd)	5.9 ± 11.6	8.9 ± 12.2	0.58

[a] in ever smokers only

8.1.2 Results

8.1.2.1 N7-Methyldeoxyguanosine (N7-medG)

The N7-medG adduct was detected in 20 out of 22 samples; all 8 current smokers and 12 / 14 current non-smokers, who were defined as former and never smokers. The mean N7-medG level was 0.41 ± 0.32 N7-medG/10^6dG with values ranging from 0.1 to 1.37 N7-medG/10^6dG, representing a thirteen fold variation between individuals in which the adduct was detected (Figure 8.1a). Levels were higher in cases compared to controls (0.55 ± 0.39 vs. 0.34 ± 0.29 N7-medG/10^6dG, p = 0.18) but this difference did not reach significance. There was no association between adduct level and smoking status, gender (Table 8.2) or age ($R^2 = 0.06$, p = 0.59).

8.1.2.2 8-Oxodeoxyguanosine (8-oxodG)

The adduct 8-oxodG was detected in all 22 DNA samples analysed. The mean level was 8.40 ± 3.58 8-oxodG/10^6dG ranging from 4.2 to 18.4 8-oxodG/10^6dG (Figure 8.1b). There was a greater than four fold variation in adduct level between individuals but no association with adduct level and smoking status, disease status, gender (Table 8.2) or age ($R^2 = 0.11$, p = 0.13).

Table 8.2 N7-medG and 8-oxodG levels in all subjects and according to smoking status, disease status and gender.

Groups Compared	Numbers Subjects	N7-medG / 10^6 dG Mean +/- SD	p value	8-oxodG / 10^6 dG Mean +/- SD	p value
All Subjects	22	0.41 ± 0.33	-	8.40 ± 3.58	-
Current Smokers	8	0.45 ± 0.27	0.72	9.35 ± 3.27	0.36
Current Non Smokers	14	0.39 ± 0.37		7.85 ± 3.74	
Cases	7	0.55 ± 0.39	0.18	7.86 ± 1.69	0.64
Controls	15	0.34 ± 0.29		8.64 ± 4.21	
Male	16	0.47 ± 0.37	0.17	8.78 ± 3.88	0.43
Female	6	0.25 ± 0.16		7.38 ± 2.59	

(a) **N7-medG**
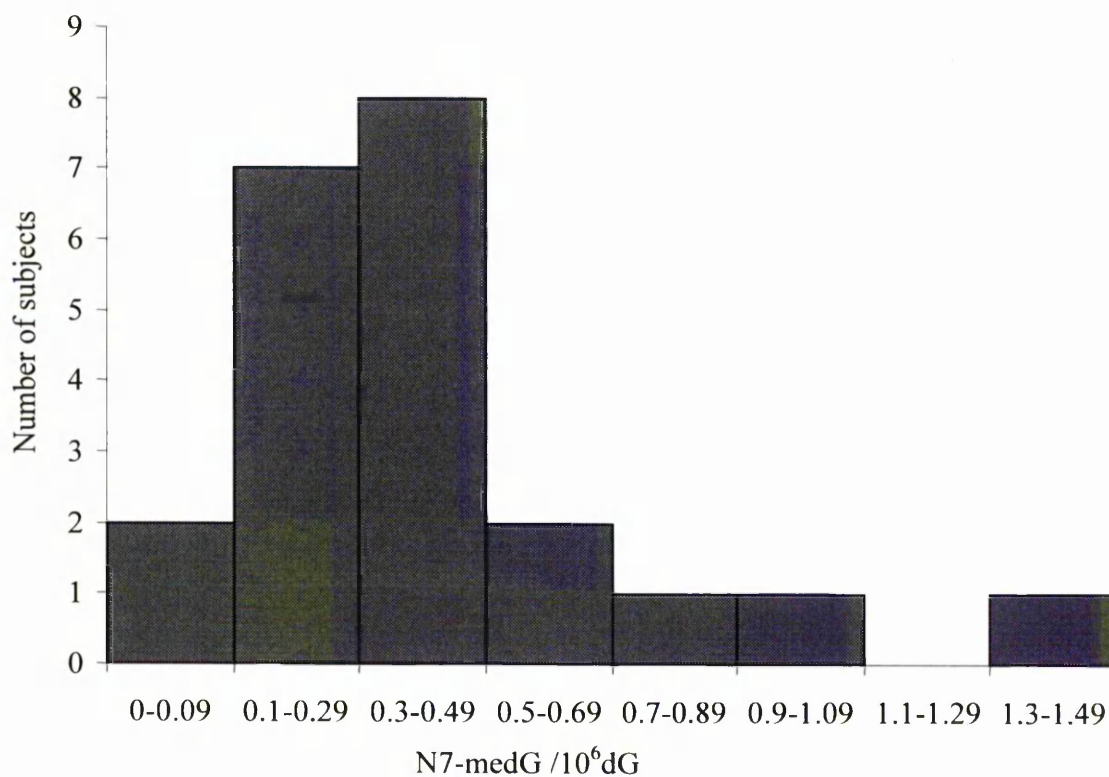
(b) **8-oxodG**
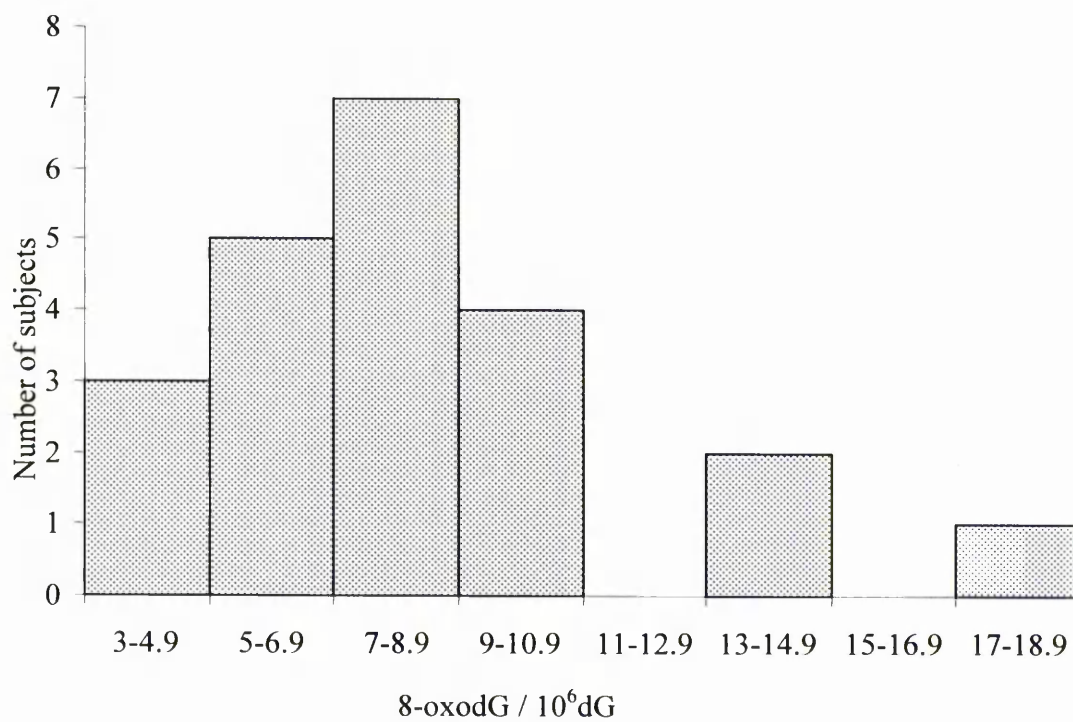
Figure 8.1 The distribution of (a) N7-medG and (b) 8-oxodG in the study population.

8.1.2.3 Comparison of N7-medG and 8-oxodG Levels

Mean 8-oxodG levels (8.40 ± 3.58 8-oxodG/10^6dG) were approximately 20 times greater than mean N7-medG levels (0.41 ± 0.33 N7-medG/10^6dG). There was no correlation between the two adducts in the 20 samples where values for both were available ($R^2 < 0.01$) (Figure 8.2).
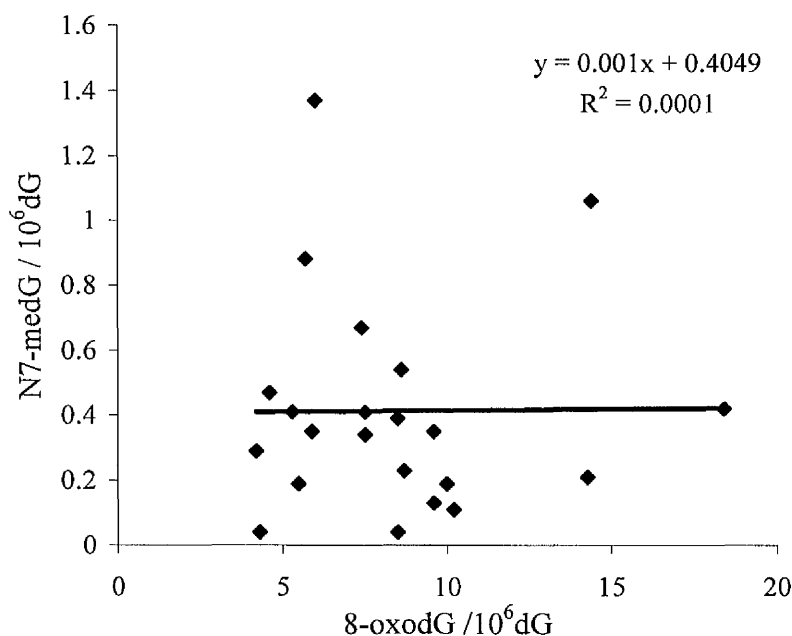
Figure 8.2 N7-medG and 8-oxodG levels in PBMCs (n = 20).

8.1.2.4 N7-medG Levels and DNA Repair Activity (MGMT and MPG)

MGMT activity was detected in all samples (n = 22) (nuclear PBMC extracts). The mean MGMT activity was 1.51 ± 0.80 fmole/μgDNA (range 0.43 to 3.45 fmole/μgDNA). MPG activity was successfully measured in 14 / 22 samples assayed. Mean MPG activity was 3.59 ± 0.65 fmole/μgDNA/hr (range 2.93 to 5.17 fmole/μgDNA/hr).

Values of N7-medG and MGMT were log transformed to normalise distribution of data. Log N7-medG levels were inversely correlated with log nuclear MGMT activity ($R^2 = 0.27$, p = 0.01) (Figure 8.3). The study was split according to smoking status and the correlation between log MGMT and log N7-medG reanalysed. In current non-smokers the inverse correlation was stronger ($R^2 = 0.56$, p = 0.002, n = 14) (Figure 8.4a). However, in current smokers there was a positive correlation, that did not reach significance ($R^2 = 0.35$, p = 0.12, n = 8) (Figure 8.4b).
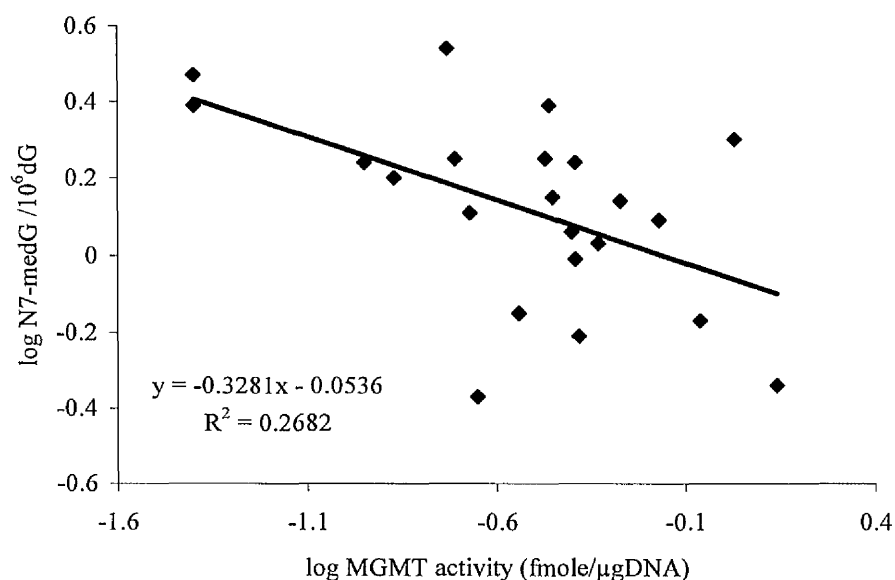
Figure 8.3 Inverse correlation between log N7-medG levels and log MGMT activity in nuclear PBMC extracts.
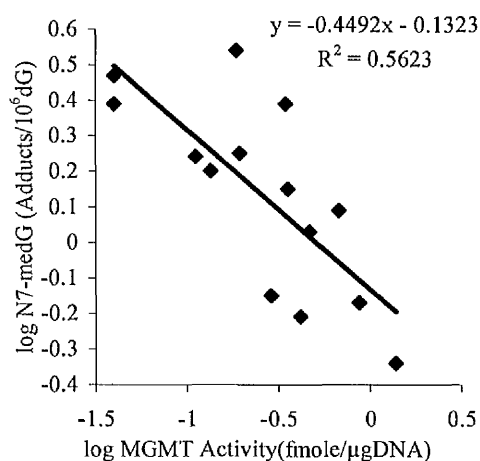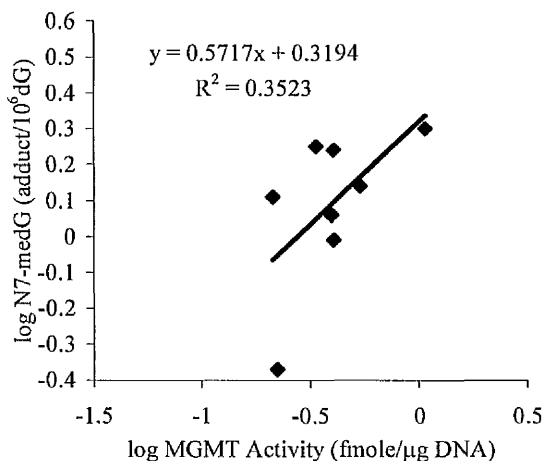
Figure 8.4 Correlation between log N7-medG levels and nuclear log MGMT activity in PBMCs from (a) current non-smokers and (b) current smokers.

There was no correlation between N7-medG and MPG activity levels ($R^2 < 0.01$, p = 0.95) (Figure 8.5), although MPG activity data was only available on 14 samples.

Figure 8.5 inside plot: y = 0.0093x + 0.4787, $R^2 = 0.0003$
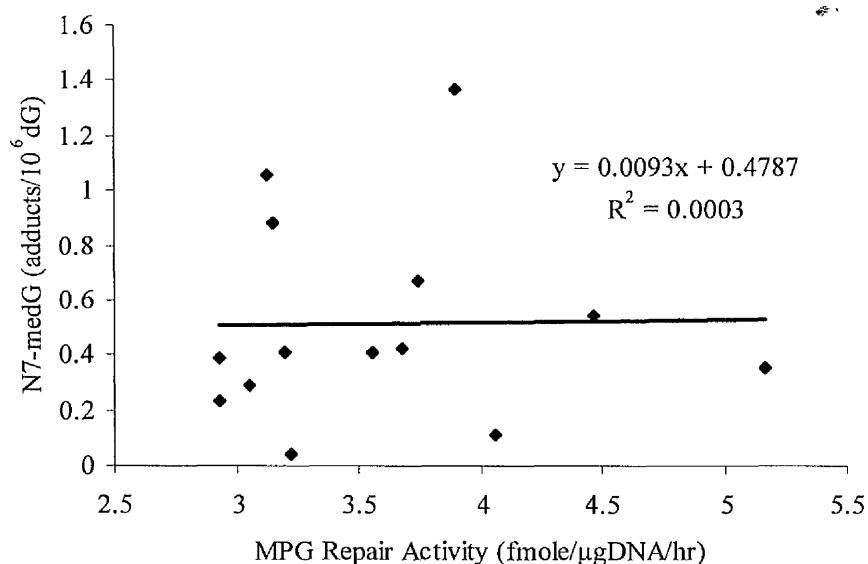
Figure 8.5 N7-medG levels and MPG activity in nuclear PBMC extracts.

8.1.3   Discussion

DNA adduct levels in PBMCs have previously been associated with risk of lung cancer (Li, Firozi et al. 2001). Elevated levels of bulky adducts in PBMCs have been shown to predict increased risk of lung cancer development after up to thirteen years of follow up (Tang, Phillips et al. 2001). Lymphocyte levels of N7-meG have also been shown to correlate with bronchial levels (Mustonen, Schoket et al. 1993). PBMCs are therefore believed to be a reasonable surrogate tissue for study of DNA adducts with respect to lung cancer.

Measurement of oxidative DNA damage can be prone to difficulties due to artefactual oxidation (Collins, Cadet et al. 2004). Two recent reports have suggested a normal 8-oxodG measurement should approximate to five 8-oxodG / 10^6dG (Gedik and Collins 2005; Poulsen 2005). The levels of 8-oxodG detected in this study are consistent with this 'normal' and also previous measurements using HPLC-ECD in human PBMCs (mean values between 1.19 – 9.57 8-oxodG / 10^6dG in non-smokers) (Lenton, Therriault et al. 1999; Thompson, Heimendinger et al. 1999; Bianchini, Elmstahl et al. 2000; Bianchini, Jaeckel et al. 2001; Gedik, Boyle et al. 2002; Gackowski, Speina et al. 2003). The levels of N7-medG in this study are lower than expected. Previous mean values ranged from 0.5–5.4 and 1.36–9.2 N7-medG/10^6dG in non-smokers and

smokers respectively in lymphocyte DNA using a [^{32}P]-postlabelling assay (Mustonen and Hemminki 1992; Kato, Petruzzelli et al. 1993; Mustonen, Schoket et al. 1993; Kumar and Hemminki 1996; Szyfter, Hemminki et al. 1996; Zhao, Tyndyk et al. 1999). This may reflect systematic differences between the immunoslotblot (ISB) and [^{32}P]-postlabelling assays; but previous work as part of method development showed that the ISB method demonstrated a good correlation with HPLC-[^{32}P]-postlabelling (Harrison, Wood et al. 2001). N7-medG levels in this sample of PBMCs may simply be slightly lower than previously reported.

The first important conclusion from this study is that levels of mutagenic DNA adducts formed as a result of exposure to reactive oxygen species are greater than those formed as a consequence of exposure to methylating agents. The mean level of 8-oxodG detected in PBMCs was 20 times greater than the mean level of N7-medG. N7-medG is not a directly mutagenic lesion but is used as a marker of methylation damage to DNA. The most mutagenic lesion caused by methylating agents is O^6-meG, which was not measured, but is generated by nitroso compounds in concentrations 10 fold lower than N7-meG. The levels of mutagenic oxidative DNA damage in the form of 8-oxodG might be 200 times greater than the mutagenic lesion, O^6-meG. If the mutagenicity of the lesions is assumed to be equal *in vivo*, then exposure to reactive oxygen species may be more important in carcinogenesis than exposure to methylating agents.

The second conclusion from this study is that levels of 8-oxodG and N7-medG are not correlated. Levels of DNA adducts represent a steady state between degree of exposure to agents that can induce adduct formation, the metabolic processing of these agents, the rate of damage formation and the rate of DNA damage repair. Differences in any or all of these factors may explain the lack of correlation seen in this study. The power of this study was too limited to draw any conclusions about any associations between smoking exposure and damage levels or other demographic variables and adduct levels. To detect a 50% increase in N7-medG and 8-oxodG levels in smokers as compared to non-smokers would have required approximately 50 and 20 subjects in each group respectively.

Both endogenous processes and exposure to exogenous agents are sources of methylating agents and reactive oxygen species. Many of these agents, for example carcinogens in tobacco smoke, require metabolic activation to achieve their carcinogenic potential; this is balanced by detoxification pathways (discussed in section 1.3.5). The balance and effectiveness of these metabolic pathways can be influenced by genotype. For example, Lewis *et al* showed that N7-

medG levels in DNA extracted from bronchial lavage cells might vary with metabolic genotype (Lewis, Cherry et al. 2004).

DNA repair processes are a further source of variability in adduct level. Both N7-medG and 8-oxodG are repaired by the base excision repair (BER) pathway (Sancar, Lindsey-Boltz et al. 2004). Lymphocytes have been shown to contain DNA repair enzymes that will remove N7-medG and 8-oxodG from DNA (Singer and Brent 1981; Paz-Elizur, Krupsky et al. 2003). There is limited evidence of correlation between different BER activities e.g. repair of 8-oxodG and uracil (Allinson, Sleeth et al. 2004) but no clear information regarding the relationship between MPG and hOGG1 activity. MPG activity was available in 14 / 22 of the samples for this study; there was no association with N7-medG levels. However, a significant inverse association was seen with MGMT activity both in the study as a whole and a stronger association in current non-smokers. Increased N7-medG levels may be found in subjects more susceptible to the damaging effects of methylating agents and this is reflected in relatively low nuclear MGMT levels. MGMT is not involved in the repair of N7-medG but previous results (Sections 5.2.2 and 6.1.4) have shown a significant correlation between MPG and MGMT activity both in whole PBMC extract and nuclear PBMC extract. The reverse trend was seen in current smokers, this may represent upregulation of nuclear MGMT levels in subjects with ongoing methylating exposure. However, the small size of the study means that care should be taken in not over-interpreting these findings; further work needs to be carried out to clarify these results.

Chapter Nine

Final Discussion

The fundamental aim of the work detailed in this thesis was to investigate factors that might modify an individual's susceptibility to tobacco smoke carcinogens and so determine the risk of lung cancer development. This was achieved by comparing individuals with chronic tobacco smoke exposure, some of whom had developed lung cancer and others who were cancer free. MPG repair activity was found to be significantly greater in cases than controls using nuclear extract from PBMCs. The present study is the first to report MPG repair activity in human samples with respect to risk of cancer. Females with above median MPG activity were found to be six times more likely to have lung cancer than those with below median activity. Measurement of MPG in female smokers might form one aspect of risk stratification into high and low risk individuals for lung cancer development. Clearly further work is required. To ascertain a causative role of elevated MPG activity in the development of lung cancer, replication of these findings in a prospective study would be required. DNA repair activity in PBMCs has previously been associated with lung cancer (Paz-Elizur, Krupsky et al. 2003). hOGG1 activity in this study was consistent over time and activity in PBMCs was strongly correlated with activity in lung tissue. Further investigation of MPG (and also MGMT) activity in human samples could be undertaken along similar lines. For instance, MPG activity could be measured in a cohort of individuals on several occasions to determine the variability of repair activity over time. Activity could also be measured in subjects before and after surgical resection of lung cancer to determine whether the presence of malignancy effects enzyme activity. The recruitment of patients from a Bronchoscopy Unit provides the opportunity of sampling bronchial epithelial cells from subjects with and without lung cancer. Activity in bronchial epithelial cells or whole lung tissue could be compared to activity in PBMCs to examine whether there was a correlation. Polymorphic variation of the MPG gene has previously been reported (Rusin, Samojedny et al. 1999), genotyping of this population would allow the investigation of whether genetic polymorphisms accounted for measured differences in repair activity. The association of high MPG activity and lung cancer was particular to female cases and controls and not to males. The cause of this gender specific effect is not known, although one study has shown a significant interaction between MPG and ERα (Likhite, Cass et al. 2004). The association with oestrogen receptors could therefore be further investigated. The role of MPG in other cancers such as breast cancer should also be investigated.

The possible role of the DNA repair protein MGMT in lung cancer development formed a significant aspect of the work in this thesis. The relationship between MGMT genotype, repair activity, allelic expression imbalance and lung cancer was investigated. Genotyping was undertaken to measure the distribution of two single nucleotide polymorphisms which previous

work had shown to be significantly associated with MGMT expression (Margison, Heighway et al. 2005). Analysis of the distribution of the codon 178 polymorphism (K178R) of the MGMT gene revealed no significant association with lung cancer; however RR homozygotes were much less commonly seen in cases than controls, this was a novel finding that warrants further investigation in a larger study. In a recent study by Margison *et al* no difference was detected in the ability of either allele to repair O^6-meG (Margison, Heighway et al. 2005). A functional difference has been reported between the K178 and 178R variant in the repair of the pyridyloxobutyl adduct O^6-pobG (Mijal, Kanugula et al. 2006). The possible reduction in lung cancer risk in RR homozygotes might therefore be due to differences in repair of pyridyloxobutyl adducts.

The second polymorphism involved a C/T transition in the first intron of the MGMT gene. Heterozygotes for this polymorphism were shown to be at lower risk of lung cancer development only when cases were compared to chest clinic controls. A significant protective effect towards squamous cell carcinoma was seen in both heterozygotes (C/T) and homozygotes (T/T) for this polymorphism. A chronic smoker who was homozygous for the T allele might have a five fold reduction in risk of squamous cell carcinoma, the most common lung cancer cell type. This was not a primary outcome of the study but discovered after subset analysis in a small group of cases, once again confirmation of this finding should be undertaken in a larger study. However, evidence taken from the present study suggests that genetic polymorphisms of the MGMT gene might modify the risk of lung cancer development, possibly through altered MGMT function or expression.

Allelic expression imbalance has been reported in two previous studies to be a common feature of MGMT expression (Heighway, Margison et al. 2003; Margison, Heighway et al. 2005). The present study for the first time measured relative allelic expression in two distinct normal tissues and also tumour tissue. The presence and extent of AEI was similar in both blood and normal lung suggesting that the cause was not environmental but genetic, possibly due to a polymorphism located in the promoter region. AEI was also seen in tumour tissue and in the small number of samples analysed, the pattern of imbalance was similar in normal tissue and tumour tissue although the degree of imbalance was stronger in tumours. This finding was not due to differences in gene copy number as the ratio of genomic DNA remained 1:1 in all samples analysed. The codon 178 polymorphism was used simply as a marker of imbalance, the cause of the imbalance might involve a polymorphism in the promoter region of the gene; screening for such a polymorphism could be undertaken, although a larger number of samples

would be required. The role of promoter hypermethylation could also be investigated as a cause of differences in allelic imbalance especially in tumour tissue, but this is not thought to be allele specific.

MGMT activity was measured in nuclear extract from PBMCs. Nuclear activity was not correlated with genotype and was not associated with lung cancer overall. However, activity was significantly higher in current smoking cases compared to controls. Nuclear activity was shown in the present study to approximate to only 5% of whole cellular activity. This was a consistent finding in both PBMCs and MCF-7 cells; but contrary to several previous studies of MGMT cellular localisation using immunohistochemistry (Ayi, Loh et al. 1992; Lee, Rafferty et al. 1992; Brent, von Wronski et al. 1993; Lim and Li 1996). Further analysis of MGMT subcellular localisation is warranted to investigate this discrepancy. The measurement and calculation of the relative amounts of active and inactivate MGMT protein in nuclear extract might help to determine whether non-specific binding of antibodies to both active and inactive protein could explain some of the observed differences between studies.

The protective effects of DNA repair pathways may be disadvantageous when treating cancer, resulting in treatment resistance to the cytotoxic effects induced by chemotherapeutic agents or radiotherapy. For example, resistance to certain methylating agents e.g. temozolomide, is facilitated through DNA repair by MGMT (Margison, Santibanez Koref et al. 2002). Inhibition of DNA repair may therefore increase the clinical effectiveness of these agents (Clemons, Kelly et al. 2005). A recent phase 1 trial of the MGMT inhibitor lomeguatrib has been undertaken in combination with temozolomide (Ranson, Middleton et al. 2006). Enhanced temozolomide activity has also been investigated using pharmacological inhibitors of PARP1 (Curtin 2005; Plummer 2006). The effectiveness of such therapy awaits further clinical trials. However, modulation of the DNA repair response is an area of development with a direct clinical application in the setting of chemotherapeutic agents. Modification of MPG activity may also be of benefit in the treatment of cancers with alkylating agents. Overexpression of MPG in a breast cancer cell line resulted in increased sensitivity to temozolomide (Rinne, Caldwell et al. 2004). Reduction in MPG activity has also been shown to increase sensitivity to alkylating agents in HeLa cell lines (Paik, Duncan et al. 2005). A recent study by Fishel *et al*, reported increased temozolomide induced ovarian cancer cell killing with overexpression of MPG (Fishel, He et al. 2007). The cell killing effect of temozolomide was significantly increased with the concurrent use of methoxyamine, a blocker of the base excision repair pathway. Gender specific differences

in repair of alkylation damage / base excision repair pathway, as identified in the present study, may be of relevance to future drug development work.

A significant and possible confounding factor in the present study was the high degree of tobacco exposure reported by study participants. The association of DNA repair activity with smoking should be investigated further in a specific study. The use of cotinine measurements would help to more accurately categorise current and former smokers. Without this external measure there is little doubt that misclassification of smoking status occurs with self reporting (Lewis, Cherry et al. 2003). One way of examining the association between smoking and repair activity would be to study a cohort of smokers attending a smoking cessation clinic. Enzyme activity could be measured before and at several time intervals after smoking cessation to determine if smoking affects enzyme activity. In addition to DNA repair activity, measurements of DNA damage, for example N7-medG and 8-oxodG, could also be taken. The temporal relationship between tobacco smoke exposure and DNA damage is not known for these adducts, although studies of anti-benzo[a]pyrene diolepoxide-DNA adduct antibodies suggest persistence for several months (Pulera, Petruzzelli et al. 1997). High levels of aromatic DNA adducts measured in PBMCs from healthy current smokers have been associated with increased risk of lung cancer in one prospective study (Tang, Phillips et al. 2001). This suggests that levels of DNA adducts may well be increased in individuals more likely to develop lung cancer and that the use of blood as a surrogate tissue is appropriate. The present work contained a pilot study examining the levels of markers of oxidative and alkylation damage; no significant difference between cases and controls was found. However, no firm conclusions could be drawn because of the small numbers involved. A larger study measuring both N7-medG and 8-oxodG in PBMCs and bronchial epithelial cells would help to determine whether DNA adducts are helpful in distinguishing case status. A prospective study would be required to ascertain whether high levels of adduct predicted lung cancer outcome.

In the present study, genetic polymorphisms in both phase 2 metabolic enzymes and DNA repair proteins were shown to affect lung cancer risk. The GSTM1 null genotype was associated with an increased risk of lung cancer in reference to bronchoscopy (OR$_{adj}$ 1.57, 95% CI 1.02-2.43) and chest clinic controls (OR$_{adj}$ 1.86, 95% CI 1.09-3.15). Real time PCR techniques were used to measure GSTM1 gene copy number and a significant dose effect was seen with respect to lung cancer risk when cases were compared to chest clinic controls ($p = 0.02$). This effect was not as convincing in reference to bronchoscopy controls ($p = 0.07$). Individuals homozygous for the GSTM1 gene represented only 7.2% of study controls; this was a significantly lower proportion

259

than heterozygote subjects (46.3%). Almost all previous association studies of GSTM1 and lung cancer did not differentiate gene copy number. The major comparison would therefore have been between null individuals and heterozygotes; the protective effects of the homozygous genotype may therefore have been underestimated. The results of this study suggest that for an accurate assessment of the relationship between GSTM1 and lung cancer risk gene copy number should be determined. A much larger study would be required to confirm the dose dependent effect given the rarity of subjects with two copies of the GSTM1 gene. The GSTT1 and hOGG1 genotypes were not associated with lung cancer. However, after stratification by GSTM1 status, lung cancer risk increased in line with GSTM1 copy number in subjects with the GSTT1 null genotype. The mechanism for this interaction was unclear. In a similar way to GSTM1 the copy number of GSTT1 would be required for a more accurate assessment of its impact on lung cancer risk and interaction with GSTM1. This was not performed in the current study but could be undertaken at a future date.

Assessment of the interaction between different 'at risk' genotypes could be extended in future work. The GSTM1 null genotype might have a greater significance in individuals who have increased phase I activation of carcinogens. For example, CYP1A1 has a central role in the metabolism of polycyclic aromatic hydrocarbons and activation of benzo[a]pyrene, an important lung carcinogen (Houlston 2000). A CYP1A1*2 polymorphism of this gene has been shown to be more active than the wildtype CYP1A1*1 enzyme and to carry a marginal increased risk of lung cancer in homozygous subjects (Bartsch, Nair et al. 2000). Individually CYP1A1*2 and GSTM1 null individuals have been shown to have increased specific benzo[a]pyrene diol epoxide (BPDE)-DNA adducts in lung tissue. The levels of this adduct were found to be highest in lung tissue of subjects with the combined CYP1A1*2/*2 and GSTM1 null genotype (Alexandrov, Cascorbi et al. 2002). Belogubova *et al* showed that the combination of the CYP1A1*2 and GSTM1 null genotype increased the risk of squamous cell lung cancer (Belogubova, Ulibina et al. 2006). It would therefore be interesting to see if there was an additive interaction between the GSTM1 null genotype and CYP1A1*2 polymorphism with respect to lung cancer risk and also analyse whether this interaction was affected by GSTM1 copy number.

A second phase I enzyme, CYP2A13 is the most efficient CYP450 enzyme involved in the metabolic activation of NNK and is also found at significant levels within the lung (Su, Bao et al. 2000). Polymorphisms that result in lower enzyme activity (e.g. Arg257Cys, Arg101stop) might also reduce lung cancer risk because of a reduction in activated NNK. A significant

reduction in risk of adenocarcinoma was reported by Wang *et al* in a Chinese population (Arg257Cys) (Wang, Tan et al. 2003); however, no association was found with the non-functioning Arg101stop CYP2A13 polymorphism and lung cancer in a French population (Cauffiez, Lo-Guidice et al. 2004). CYP2A13 genotyping in the current study could be performed and compared to both lung cancer outcome and also to measures of alkylation damage e.g. N7-meG or measures of MGMT and MPG repair activity, both enzymes that are involved in repair of alkylation damage.

The disadvantage of this study with respect to measuring the effect of such genotypes on lung cancer risk is the number of study participants. The study was powerful enough to detect differences in distribution of the GSTM1 genotype, which affects 50% of the study population. However polymorphisms of the CYP1A1 and CYP2A13 gene generally affect less than 5% of Caucasian populations (Houlston 2000; Cauffiez, Lo-Guidice et al. 2004), the size of the study population required to detect significant differences in these genotypes would therefore need to be greatly increased. In principle, however, the investigation of multiple genotypes involved in the processing pathway of particular carcinogens or the repair of particular types of DNA damage is appropriate. It is likely that a combination of low penetrance genes rather than a single high risk 'lung cancer' gene will be responsible for a significant amount of the genetic modification of lung cancer risk. The identification of susceptible individuals / populations through the use of biological markers might enable smokers to be more accurately risk stratified and could result in focused screening or smoking cessation strategies. In the future, chemopreventive agents may also be available to those deemed at risk of lung cancer occurrence or recurrence.

References

(1995). "Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. Non-small Cell Lung Cancer Collaborative Group." Bmj 311(7010): 899-909.

Adhikari, S., J. A. Toretsky, L. Yuan and R. Roy (2006). "Magnesium, essential for base excision repair enzymes, inhibits substrate binding of N-methylpurine-DNA glycosylase." J Biol Chem 281(40): 29525-32.

Agra, Y., M. Pelayo, M. Sacristan, A. Sacristan, C. Serra and X. Bonfill (2003). "Chemotherapy versus best supportive care for extensive small cell lung cancer." Cochrane Database Syst Rev(4): CD001990.

Ahrendt, S. A., P. A. Decker, E. A. Alawi, Y. R. Zhu Yr, M. Sanchez-Cespedes, S. C. Yang, G. B. Haasler, A. Kajdacsy-Balla, M. J. Demeure and D. Sidransky (2001). "Cigarette smoking is strongly associated with mutation of the K-ras gene in patients with primary adenocarcinoma of the lung." Cancer 92(6): 1525-30.

Alberg, A. J. and J. M. Samet (2003). "Epidemiology of lung cancer." Chest 123(1 Suppl): 21S-49S.

Alexandrov, K., I. Cascorbi, M. Rojas, G. Bouvier, E. Kriek and H. Bartsch (2002). "CYP1A1 and GSTM1 genotypes affect benzo[a]pyrene DNA adducts in smokers' lung: comparison with aromatic/hydrophobic adduct formation." Carcinogenesis 23(12): 1969-77.

Ali-Osman, F., O. Akande, G. Antoun, J. X. Mao and J. Buolamwini (1997). "Molecular cloning, characterization, and expression in Escherichia coli of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Evidence for differential catalytic activity of the encoded proteins." J Biol Chem 272(15): 10004-12.

Allan, J. M., B. P. Engelward, A. J. Dreslin, M. D. Wyatt, M. Tomasz and L. D. Samson (1998). "Mammalian 3-methyladenine DNA glycosylase protects against the toxicity and clastogenicity of certain chemotherapeutic DNA cross-linking agents." Cancer Res 58(17): 3965-73.

Allinson, S. L., K. M. Sleeth, G. E. Matthewman and G. L. Dianov (2004). "Orchestration of base excision repair by controlling the rates of enzymatic activities." DNA Repair (Amst) 3(1): 23-31.

Anttila, S., A. Hirvonen, H. Vainio, K. Husgafvel-Pursiainen, J. D. Hayes and B. Ketterer (1993). "Immunohistochemical localization of glutathione S-transferases in human lung." Cancer Res 53(23): 5643-8.

Arai, T., V. P. Kelly, O. Minowa, T. Noda and S. Nishimura (2006). "The study using wild-type and Ogg1 knockout mice exposed to potassium bromate shows no tumor induction despite an extensive accumulation of 8-hydroxyguanine in kidney DNA." Toxicology 221(2-3): 179-86.

Ariyoshi, N., M. Miyamoto, Y. Umetsu, H. Kunitoh, H. Dosaka-Akita, Y. Sawamura, J. Yokota, N. Nemoto, K. Sato and T. Kamataki (2002). "Genetic polymorphism of CYP2A6 gene and tobacco-induced lung cancer risk in male smokers." Cancer Epidemiol Biomarkers Prev 11(9): 890-4.

Asami, S., T. Hirano, R. Yamaguchi, Y. Tomioka, H. Itoh and H. Kasai (1996). "Increase of a type of oxidative DNA damage, 8-hydroxyguanine, and its repair activity in human leukocytes by cigarette smoking." Cancer Res **56**(11): 2546-9.

Asami, S., H. Manabe, J. Miyake, Y. Tsurudome, T. Hirano, R. Yamaguchi, H. Itoh and H. Kasai (1997). "Cigarette smoking induces an increase in oxidative DNA damage, 8-hydroxydeoxyguanosine, in a central site of the human lung." Carcinogenesis **18**(9): 1763-6.

Austin, J. H., G. D. Pearson and B. Thomashow (2002). "CT screening for early stage lung cancer." Chest **121**(5): 1725-6.

Austin, J. H., B. M. Romney and L. S. Goldsmith (1992). "Missed bronchogenic carcinoma: radiographic findings in 27 patients with a potentially resectable lesion evident in retrospect." Radiology **182**(1): 115-22.

Ayi, T. C., K. C. Loh, R. B. Ali and B. F. Li (1992). "Intracellular localization of human DNA repair enzyme methylguanine-DNA methyltransferase by antibodies and its importance." Cancer Res **52**(23): 6423-30.

Bailey-Wilson, J. E., C. I. Amos, S. M. Pinney, G. M. Petersen, M. de Andrade, J. S. Wiest, P. Fain, A. G. Schwartz, M. You, W. Franklin, C. Klein, A. Gazdar, H. Rothschild, D. Mandal, T. Coons, J. Slusser, J. Lee, C. Gaba, E. Kupert, A. Perez, X. Zhou, D. Zeng, Q. Liu, Q. Zhang, D. Seminara, J. Minna and M. W. Anderson (2004). "A major lung cancer susceptibility locus maps to chromosome 6q23-25." Am J Hum Genet **75**(3): 460-74.

Bain, C., D. Feskanich, F. E. Speizer, M. Thun, E. Hertzmark, B. A. Rosner and G. A. Colditz (2004). "Lung cancer rates in men and women with comparable histories of smoking." J Natl Cancer Inst **96**(11): 826-34.

Barbin, A., R. Wang, P. J. O'Connor and R. H. Elder (2003). "Increased formation and persistence of 1,N(6)-ethenoadenine in DNA is not associated with higher susceptibility to carcinogenesis in alkylpurine-DNA-N-glycosylase knockout mice treated with vinyl carbamate." Cancer Res **63**(22): 7699-703.

Barnes, D. E. and T. Lindahl (2004). "Repair and genetic consequences of endogenous DNA base damage in mammalian cells." Annu Rev Genet **38**: 445-76.

Barnes, P. J. (2000). "Chronic obstructive pulmonary disease." N Engl J Med **343**(4): 269-80.

Bartsch, H., U. Nair, A. Risch, M. Rojas, H. Wikman and K. Alexandrov (2000). "Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers." Cancer Epidemiol Biomarkers Prev **9**(1): 3-28.

Bastarrika, G., M. J. Garcia-Velloso, M. D. Lozano, U. Montes, W. Torre, N. Spiteri, A. Campo, L. Seijo, A. B. Alcaide, J. Pueyo, D. Cano, I. Vivas, O. Cosin, P. Dominguez, P. Serra, J. A. Richter, L. Montuenga and J. J. Zulueta (2005). "Early lung cancer detection using spiral computed tomography and positron emission tomography." Am J Respir Crit Care Med **171**(12): 1378-83.

Beckles, M. A., S. G. Spiro, G. L. Colice and R. M. Rudd (2003). "Initial evaluation of the patient with lung cancer: symptoms, signs, laboratory tests, and paraneoplastic syndromes." Chest **123**(1 Suppl): 97S-104S.

Belanich, M., T. Randall, M. A. Pastor, J. T. Kibitel, L. G. Alas, M. E. Dolan, S. C. Schold, Jr., M. Gander, F. J. Lejeune, B. F. Li, A. B. White, P. Wasserman, M. L. Citron and D. B. Yarosh (1996). "Intracellular Localization and intercellular heterogeneity of the human DNA repair protein O(6)-methylguanine-DNA methyltransferase." Cancer Chemother Pharmacol 37(6): 547-55.

Belinsky, S. A., J. F. Foley, C. M. White, M. W. Anderson and R. R. Maronpot (1990). "Dose-response relationship between O6-methylguanine formation in Clara cells and induction of pulmonary neoplasia in the rat by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone." Cancer Res 50(12): 3772-80.

Belogubova, E. V., A. V. Togo, M. B. Karpova, E. Kuligina, K. G. Buslova, J. M. Ulibina, V. G. Lemehov, S. M. Romanenko, V. A. Shutkin, K. P. Hanson, A. Hirvonen and E. N. Imyanitov (2004). "A novel approach for assessment of cancer predisposing roles of GSTM1 and GSTT1 genes: use of putatively cancer resistant elderly tumor-free smokers as the referents." Lung Cancer 43(3): 259-66.

Belogubova, E. V., Y. M. Ulibina, I. K. Suvorova, E. Kuligina, M. B. Karpova, V. A. Shutkin, A. V. Koloskov, A. P. Kuchinskiy, A. V. Togo, K. P. Hanson, A. Hirvonen and E. N. Imyanitov (2006). "Combined CYP1A1/GSTM1 at-risk genotypes are overrepresented in squamous cell lung carcinoma patients but underrepresented in elderly tumor-free subjects." J Cancer Res Clin Oncol 132(5): 327-31.

Benhamou, S., W. J. Lee, A. K. Alexandrie, P. Boffetta, C. Bouchardy, D. Butkiewicz, J. Brockmoller, M. L. Clapper, A. Daly, V. Dolzan, J. Ford, L. Gaspari, A. Haugen, A. Hirvonen, K. Husgafvel-Pursiainen, M. Ingelman-Sundberg, I. Kalina, M. Kihara, P. Kremers, L. Le Marchand, S. J. London, V. Nazar-Stewart, M. Onon-Kihara, A. Rannug, M. Romkes, D. Ryberg, J. Seidegard, P. Shields, R. C. Strange, I. Stucker, J. To-Figueras, P. Brennan and E. Taioli (2002). "Meta- and pooled analyses of the effects of glutathione S-transferase M1 polymorphisms and smoking on lung cancer risk." Carcinogenesis 23(8): 1343-50.

Beranek, D. T. (1990). "Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents." Mutat Res 231(1): 11-30.

Berdal, K. G., R. F. Johansen and E. Seeberg (1998). "Release of normal bases from intact DNA by a native DNA repair enzyme." Embo J 17(2): 363-7.

Bianchini, F., S. Elmstahl, C. Martinez-Garcia, A. L. van Kappel, T. Douki, J. Cadet, H. Ohshima, E. Riboli and R. Kaaks (2000). "Oxidative DNA damage in human lymphocytes: correlations with plasma levels of alpha-tocopherol and carotenoids." Carcinogenesis 21(2): 321-4.

Bianchini, F., A. Jaeckel, P. Vineis, C. Martinez-Garcia, S. Elmstahl, A. L. van Kappel, H. Boeing, H. Ohshima, E. Riboli and R. Kaaks (2001). "Inverse correlation between alcohol consumption and lymphocyte levels of 8-hydroxydeoxyguanosine in humans." Carcinogenesis 22(6): 885-90.

Blomeke, B., M. J. Greenblatt, V. D. Doan, E. D. Bowman, S. E. Murphy, C. C. Chen, S. Kato and P. G. Shields (1996). "Distribution of 7-alkyl-2'-deoxyguanosine adduct levels in human lung." Carcinogenesis 17(4): 741-8.

Boffetta, P., F. Nyberg, A. Mukeria, S. Benhamou, V. Constantinescu, H. Batura-Gabryel, I. Bruske-Hohlfeld, G. Schmid, L. Simonato, P. Pelkonen and J. Hall (2002). "O6-Alkylguanine-

DNA-alkyltransferase activity in peripheral leukocytes, smoking and risk of lung cancer." Cancer Lett **180**(1): 33-9.

Boiteux, S. and J. P. Radicella (2000). "The human OGG1 gene: structure, functions, and its implication in the process of carcinogenesis." Arch Biochem Biophys **377**(1): 1-8.

Bouziane, M., F. Miao, S. E. Bates, L. Somsouk, B. C. Sang, M. Denissenko and T. R. O'Connor (2000). "Promoter structure and cell cycle dependent expression of the human methylpurine-DNA glycosylase gene." Mutat Res **461**(1): 15-29.

Boysen, G. and S. S. Hecht (2003). "Analysis of DNA and protein adducts of benzo[a]pyrene in human tissues using structure-specific methods." Mutat Res **543**(1): 17-30.

Brabender, J., H. Usadel, R. Metzger, P. M. Schneider, J. Park, D. Salonga, D. D. Tsao-Wei, S. Groshen, R. V. Lord, N. Takebe, S. Schneider, A. H. Holscher, K. D. Danenberg and P. V. Danenberg (2003). "Quantitative O(6)-methylguanine DNA methyltransferase methylation analysis in curatively resected non-small cell lung cancer: associations with clinical outcome." Clin Cancer Res **9**(1): 223-7.

Brambilla, E. and B. Corrin (2003). Aetiology and pathogenesis of carcinoma of the lung. Respiratory Medicine. J. G. Gibson, D. M. Geddes, U. Costabel, P. J. Sterk and B. Corrin, Saunders. **2:** 1819-1827.

Brasch-Andersen, C., L. Christiansen, Q. Tan, A. Haagerup, J. Vestbo and T. A. Kruse (2004). "Possible gene dosage effect of glutathione-S-transferases on atopic asthma: using real-time PCR for quantification of GSTM1 and GSTT1 gene copy numbers." Hum Mutat **24**(3): 208-14.

Bray, N. J., P. R. Buckland, M. J. Owen and M. C. O'Donovan (2003). "Cis-acting variation in the expression of a high proportion of genes in human brain." Hum Genet **113**(2): 149-53.

Brennan, P., C. C. Hsu, N. Moullan, N. Szeszenia-Dabrowska, J. Lissowska, D. Zaridze, P. Rudnai, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, F. Gemignani, A. Chabrier, J. Hall, R. J. Hung, P. Boffetta and F. Canzian (2005). "Effect of cruciferous vegetables on lung cancer in patients stratified by genetic status: a mendelian randomisation approach." Lancet **366**(9496): 1558-60.

Brent, T. P., M. A. von Wronski, C. C. Edwards, M. Bromley, G. P. Margison, J. A. Rafferty, C. N. Pegram and D. D. Bigner (1993). "Identification of nitrosourea-resistant human rhabdomyosarcomas by in situ immunostaining of O6-methylguanine-DNA methyltransferase." Oncol Res **5**(2): 83-6.

Burns, D. M. (1994). Tobacco Smoking. Epidemiology of Lung Cancer. J. M. Samet, Marcel Dekker. **74:** 15-46.

Calikoglu, M., L. Tamer, N. Ates Aras, S. Karakas and B. Ercan (2006). "The Association Between Polymorphic Genotypes of Glutathione S-Transferases and COPD in the Turkish Population." Biochem Genet.

Calleja, F., J. G. Jansen, H. Vrieling, F. Laval and A. A. van Zeeland (1999). "Modulation of the toxic and mutagenic effects induced by methyl methanesulfonate in Chinese hamster ovary cells by overexpression of the rat N-alkylpurine-DNA glycosylase." Mutat Res **425**(2): 185-94.

Capewell, S., R. Sankaran, D. Lamb, M. McIntyre and M. F. Sudlow (1991). "Lung cancer in lifelong non-smokers. Edinburgh Lung Cancer Group." Thorax **46**(8): 565-8.

Cappelli, E., T. Hazra, J. W. Hill, G. Slupphaug, M. Bogliolo and G. Frosina (2001). "Rates of base excision repair are not solely dependent on levels of initiating enzymes." Carcinogenesis 22(3): 387-93.

Carbone, P. (1970). "Lung cancer: perspectives and prospects." Ann Intern Med 73(6): 1003-24.

Cauffiez, C., J. M. Lo-Guidice, S. Quaranta, D. Allorge, D. Chevalier, S. Cenee, R. Hamdan, M. Lhermitte, J. J. Lafitte, C. Libersa, J. F. Colombel, I. Stucker and F. Broly (2004). "Genetic polymorphism of the human cytochrome CYP2A13 in a French population: implication in lung cancer susceptibility." Biochem Biophys Res Commun 317(2): 662-9.

Cerda, S. R., S. S. Chu, P. Garcia, J. Chung, J. D. Grumet, B. Thimmapaya and S. A. Weitzman (1999). "Regulation of expression of N-methylpurine DNA glycosylase in human mammary epithelial cells: role of transcription factor AP-2." Chem Res Toxicol 12(11): 1098-109.

Cerda, S. R., P. W. Turk, A. D. Thor and S. A. Weitzman (1998). "Altered expression of the DNA repair protein, N-methylpurine-DNA glycosylase (MPG) in breast cancer." FEBS Lett 431(1): 12-8.

Chan, A. O., S. G. Kim, A. Bedeir, J. P. Issa, S. R. Hamilton and A. Rashid (2003). "CpG island methylation in carcinoid and pancreatic endocrine tumors." Oncogene 22(6): 924-34.

Chen, L., A. Elahi, J. Pow-Sang, P. Lazarus and J. Park (2003). "Association between polymorphism of human oxoguanine glycosylase 1 and risk of prostate cancer." J Urol 170(6 Pt 1): 2471-4.

Cheng, K. C., D. S. Cahill, H. Kasai, S. Nishimura and L. A. Loeb (1992). "8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions." J Biol Chem 267(1): 166-72.

Cheng, S. L., C. J. Yu, C. J. Chen and P. C. Yang (2004). "Genetic polymorphism of epoxide hydrolase and glutathione S-transferase in COPD." Eur Respir J 23(6): 818-24.

Chung, F. L., H. J. Chen and R. G. Nath (1996). "Lipid peroxidation as a potential endogenous source for the formation of exocyclic DNA adducts." Carcinogenesis 17(10): 2105-11.

Church, D. F. and W. A. Pryor (1985). "Free-radical chemistry of cigarette smoke and its toxicological implications." Environ Health Perspect 64: 111-26.

Chute, J. P., T. Chen, E. Feigal, R. Simon and B. E. Johnson (1999). "Twenty years of phase III trials for patients with extensive-stage small-cell lung cancer: perceptible progress." J Clin Oncol 17(6): 1794-801.

Chyczewski, L., J. Niklinski, E. Chyczewska, W. Niklinska and W. Naumnik (2001). "Morphological aspects of carcinogenesis in the lung." Lung Cancer 34 Suppl 2: S17-25.

Citron, M., M. Schoenhaus, M. Graver, M. Hoffman, M. Lewis, P. Wasserman, M. Niederland, L. Kahn, A. White and D. Yarosh (1993). "O6-methylguanine-DNA methyltransferase in human normal and malignant lung tissues." Cancer Invest 11(3): 258-63.

Clemons, M., J. Kelly, A. J. Watson, A. Howell, R. S. McElhinney, T. B. McMurry and G. P. Margison (2005). "O6-(4-bromothenyl)guanine reverses temozolomide resistance in human breast tumour MCF-7 cells and xenografts." Br J Cancer 93(10): 1152-6.

Cloutier, J. F., R. Drouin, M. Weinfeld, T. R. O'Connor and A. Castonguay (2001). "Characterization and mapping of DNA damage induced by reactive metabolites of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) at nucleotide resolution in human genomic DNA." J Mol Biol 313(3): 539-57.

Cohet, C., S. Borel, F. Nyberg, A. Mukeria, I. Bruske-Hohlfeld, V. Constantinescu, S. Benhamou, P. Brennan, J. Hall and P. Boffetta (2004). "Exon 5 polymorphisms in the O6-alkylguanine DNA alkyltransferase gene and lung cancer risk in non-smokers exposed to second-hand smoke." Cancer Epidemiol Biomarkers Prev 13(2): 320-3.

Coleman, M. P., B. Rachet, L. M. Woods, E. Mitry, M. Riga, N. Cooper, M. J. Quinn, H. Brenner and J. Esteve (2004). "Trends and socioeconomic inequalities in cancer survival in England and Wales up to 2001." Br J Cancer 90(7): 1367-73.

Collins, A. R., J. Cadet, L. Moller, H. E. Poulsen and J. Vina (2004). "Are we sure we know how to measure 8-oxo-7,8-dihydroguanine in DNA from human cells?" Arch Biochem Biophys 423(1): 57-65.

Connor, E. E., J. J. Wilson and M. D. Wyatt (2005). "Effects of substrate specificity on initiating the base excision repair of N-methylpurines by variant human 3-methyladenine DNA glycosylases." Chem Res Toxicol 18(1): 87-94.

Cooke, M. S., M. D. Evans, M. Dizdaroglu and J. Lunec (2003). "Oxidative DNA damage: mechanisms, mutation, and disease." Faseb J 17(10): 1195-214.

Coombes, R. C., E. Hall, L. J. Gibson, R. Paridaens, J. Jassem, T. Delozier, S. E. Jones, I. Alvarez, G. Bertelli, O. Ortmann, A. S. Coates, E. Bajetta, D. Dodwell, R. E. Coleman, L. J. Fallowfield, E. Mickiewicz, J. Andersen, P. E. Lonning, G. Cocconi, A. Stewart, N. Stuart, C. F. Snowdon, M. Carpentieri, G. Massimini, J. M. Bliss and C. van de Velde (2004). "A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer." N Engl J Med 350(11): 1081-92.

Copas, J. B. and J. Q. Shi (2000). "Reanalysis of epidemiological evidence on lung cancer and passive smoking." Bmj 320(7232): 417-8.

Coquerelle, T., J. Dosch and B. Kaina (1995). "Overexpression of N-methylpurine-DNA glycosylase in Chinese hamster ovary cells renders them more sensitive to the production of chromosomal aberrations by methylating agents--a case of imbalanced DNA repair." Mutat Res 336(1): 9-17.

Coultas, D. B. and J. M. Samet (1992). "Occupational lung cancer." Clin Chest Med 13(2): 341-54.

CRUK (2007). "Lung cancer statistics." Cancer Research UK: http://info.cancerresearchuk.org/cancerstats/types/lung/.

Curtin, N. J. (2005). "PARP inhibitors for cancer therapy." Expert Rev Mol Med 7(4): 1-20.

Danam, R. P., X. C. Qian, S. R. Howell and T. P. Brent (1999). "Methylation of selected CpGs in the human O6-methylguanine-DNA methyltransferase promoter region as a marker of gene silencing." Mol Carcinog 24(2): 85-9.

Darby, S., D. Hill and R. Doll (2001). "Radon: a likely carcinogen at all exposures." Ann Oncol 12(10): 1341-51.

Darnton, A. J., D. M. McElvenny and J. T. Hodgson (2006). "Estimating the Number of Asbestos-related Lung Cancer Deaths in Great Britain from 1980 to 2000." Ann Occup Hyg 50(1): 29-38.

Deleyiannis, F. W. and D. B. Thomas (1997). "Risk of lung cancer among patients with head and neck cancer." Otolaryngol Head Neck Surg 116(6 Pt 1): 630-6.

Demple, B., T. Herman and D. S. Chen (1991). "Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes." Proc Natl Acad Sci U S A 88(24): 11450-4.

Deng, C., D. Xie, H. Capasso, Y. Zhao, L. D. Wang and J. Y. Hong (1999). "Genetic polymorphism of human O6-alkylguanine-DNA alkyltransferase: identification of a missense variation in the active site region." Pharmacogenetics 9(1): 81-7.

Denissenko, M. F., A. Pao, M. Tang and G. P. Pfeifer (1996). "Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53." Science 274(5286): 430-2.

Dianov, G. L., K. M. Sleeth, Dianova, II and S. L. Allinson (2003). "Repair of abasic sites in DNA." Mutat Res 531(1-2): 157-63.

Dockery, D. W., C. A. Pope, 3rd, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris, Jr. and F. E. Speizer (1993). "An association between air pollution and mortality in six U.S. cities." N Engl J Med 329(24): 1753-9.

Doll, R. and A. B. Hill (1950). "Smoking and carcinoma of the lung; preliminary report." Br Med J 2(4682): 739-48.

Doll, R. and A. B. Hill (1952). "A study of the aetiology of carcinoma of the lung." Br Med J 2(4797): 1271-86.

Doll, R., R. Peto, J. Boreham and I. Sutherland (2004). "Mortality in relation to smoking: 50 years' observations on male British doctors." Bmj 328(7455): 1519.

Dominioni, L., A. Imperatori, F. Rovera, A. Ochetti, G. Torrigiotti and M. Paolucci (2000). "Stage I nonsmall cell lung carcinoma: analysis of survival and implications for screening." Cancer 89(11 Suppl): 2334-44.

Drablos, F., E. Feyzi, P. A. Aas, C. B. Vaagbo, B. Kavli, M. S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug and H. E. Krokan (2004). "Alkylation damage in DNA and RNA--repair mechanisms and medical significance." DNA Repair (Amst) 3(11): 1389-407.

Dresler, C. M., C. Fratelli, J. Babb, L. Everley, A. A. Evans and M. L. Clapper (2000). "Gender differences in genetic susceptibility for lung cancer." Lung Cancer 30(3): 153-60.

Drin, I., B. Schoket, S. Kostic and I. Vincze (1994). "Smoking-related increase in O6-alkylguanine-DNA alkyltransferase activity in human lung tissue." Carcinogenesis 15(8): 1535-9.

Dubey, S., J. M. Siegfried and A. M. Traynor (2006). "Non-small-cell lung cancer and breast carcinoma: chemotherapy and beyond." Lancet Oncol 7(5): 416-24.

Eaton, D. L. and T. K. Bammler (1999). "Concise review of the glutathione S-transferases and their significance to toxicology." Toxicol Sci 49(2): 156-64.

el Ghissassi, F., A. Barbin, J. Nair and H. Bartsch (1995). "Formation of 1,N6-ethenoadenine and 3,N4-ethenocytosine by lipid peroxidation products and nucleic acid bases." Chem Res Toxicol **8**(2): 278-83.

Elder, R. H., J. G. Jansen, R. J. Weeks, M. A. Willington, B. Deans, A. J. Watson, K. J. Mynett, J. A. Bailey, D. P. Cooper, J. A. Rafferty, M. C. Heeran, S. W. Wijnhoven, A. A. van Zeeland and G. P. Margison (1998). "Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate." Mol Cell Biol **18**(10): 5828-37.

Ellis, J. R. and F. V. Gleeson (2001). "Lung cancer screening." Br J Radiol **74**(882): 478-85.

el-Torky, M., F. el-Zeky and J. C. Hall (1990). "Significant changes in the distribution of histologic types of lung cancer. A review of 4928 cases." Cancer **65**(10): 2361-7.

Engelward, B. P., J. M. Allan, A. J. Dreslin, J. D. Kelly, M. M. Wu, B. Gold and L. D. Samson (1998). "A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome." J Biol Chem **273**(9): 5412-8.

Engelward, B. P., A. Dreslin, J. Christensen, D. Huszar, C. Kurahara and L. Samson (1996). "Repair-deficient 3-methyladenine DNA glycosylase homozygous mutant mouse cells have increased sensitivity to alkylation-induced chromosome damage and cell killing." Embo J **15**(4): 945-52.

Engelward, B. P., G. Weeda, M. D. Wyatt, J. L. Broekhof, J. de Wit, I. Donker, J. M. Allan, B. Gold, J. H. Hoeijmakers and L. D. Samson (1997). "Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase." Proc Natl Acad Sci U S A **94**(24): 13087-92.

Erzberger, J. P. and D. M. Wilson, 3rd (1999). "The role of Mg2+ and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis." J Mol Biol **290**(2): 447-57.

Esteller, M., S. R. Hamilton, P. C. Burger, S. B. Baylin and J. G. Herman (1999). "Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia." Cancer Res **59**(4): 793-7.

Esteller, M., M. Sanchez-Cespedes, R. Rosell, D. Sidransky, S. B. Baylin and J. G. Herman (1999). "Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients." Cancer Res **59**(1): 67-70.

Feskanich, D., R. G. Ziegler, D. S. Michaud, E. L. Giovannucci, F. E. Speizer, W. C. Willett and G. A. Colditz (2000). "Prospective study of fruit and vegetable consumption and risk of lung cancer among men and women." J Natl Cancer Inst **92**(22): 1812-23.

Fishel, M. L., Y. He, M. L. Smith and M. R. Kelley (2007). "Manipulation of base excision repair to sensitize ovarian cancer cells to alkylating agent temozolomide." Clin Cancer Res **13**(1): 260-7.

Fishel, M. L., Y. R. Seo, M. L. Smith and M. R. Kelley (2003). "Imbalancing the DNA base excision repair pathway in the mitochondria; targeting and overexpressing N-methylpurine DNA glycosylase in mitochondria leads to enhanced cell killing." Cancer Res **63**(3): 608-15.

Flanders, W. D., C. A. Lally, B. P. Zhu, S. J. Henley and M. J. Thun (2003). "Lung cancer mortality in relation to age, duration of smoking, and daily cigarette consumption: results from Cancer Prevention Study II." Cancer Res **63**(19): 6556-62.

Flehinger, B. J., M. Kimmel and M. R. Melamed (1992). "The effect of surgical treatment on survival from early lung cancer. Implications for screening." Chest **101**(4): 1013-8.

Fletcher, F., R. N. Johnston and P. Stradling (1976). "The normal chest radiograph in bronchial carcinoma." Br Med J **2**(6032): 403.

Fontana, R. S., D. R. Sanderson, L. B. Woolner, W. F. Taylor, W. E. Miller and J. R. Muhm (1986). "Lung cancer screening: the Mayo program." J Occup Med **28**(8): 746-50.

Fontana, R. S., D. R. Sanderson, L. B. Woolner, W. F. Taylor, W. E. Miller, J. R. Muhm, P. E. Bernatz, W. S. Payne, P. C. Pairolero and E. J. Bergstralh (1991). "Screening for lung cancer. A critique of the Mayo Lung Project." Cancer **67**(4 Suppl): 1155-64.

Fraga, C. G., M. K. Shigenaga, J. W. Park, P. Degan and B. N. Ames (1990). "Oxidative damage to DNA during aging: 8-hydroxy-2'-deoxyguanosine in rat organ DNA and urine." Proc Natl Acad Sci U S A **87**(12): 4533-7.

Gackowski, D., E. Speina, M. Zielinska, J. Kowalewski, R. Rozalski, A. Siomek, T. Paciorek, B. Tudek and R. Olinski (2003). "Products of oxidative DNA damage and repair as possible biomarkers of susceptibility to lung cancer." Cancer Res **63**(16): 4899-902.

Gajra, A., N. Newman, G. P. Gamble, N. Z. Abraham, L. J. Kohman and S. L. Graziano (2003). "Impact of tumor size on survival in stage IA non-small cell lung cancer: a case for subdividing stage IA disease." Lung Cancer **42**(1): 51-7.

Ganti, A. K., A. E. Sahmoun, A. W. Panwalkar, K. K. Tendulkar and A. Potti (2006). "Hormone replacement therapy is associated with decreased survival in women with lung cancer." J Clin Oncol **24**(1): 59-63.

Gao, W. M., H. H. Mady, G. Y. Yu, J. M. Siegfried, J. D. Luketich, M. F. Melhem and P. Keohavong (2003). "Comparison of p53 mutations between adenocarcinoma and squamous cell carcinoma of the lung: unique spectra involving G to A transitions and G to T transversions in both histologic types." Lung Cancer **40**(2): 141-50.

Garte, S., L. Gaspari, A. K. Alexandrie, C. Ambrosone, H. Autrup, J. L. Autrup, H. Baranova, L. Bathum, S. Benhamou, P. Boffetta, C. Bouchardy, K. Breskvar, J. Brockmoller, I. Cascorbi, M. L. Clapper, C. Coutelle, A. Daly, M. Dell'Omo, V. Dolzan, C. M. Dresler, A. Fryer, A. Haugen, D. W. Hein, A. Hildesheim, A. Hirvonen, L. L. Hsieh, M. Ingelman-Sundberg, I. Kalina, D. Kang, M. Kihara, C. Kiyohara, P. Kremers, P. Lazarus, L. Le Marchand, M. C. Lechner, E. M. van Lieshout, S. London, J. J. Manni, C. M. Maugard, S. Morita, V. Nazar-Stewart, K. Noda, Y. Oda, F. F. Parl, R. Pastorelli, I. Persson, W. H. Peters, A. Rannug, T. Rebbeck, A. Risch, L. Roelandt, M. Romkes, D. Ryberg, J. Salagovic, B. Schoket, J. Seidegard, P. G. Shields, E. Sim, D. Sinnet, R. C. Strange, I. Stucker, H. Sugimura, J. To-Figueras, P. Vineis, M. C. Yu and E. Taioli (2001). "Metabolic gene polymorphism frequencies in control populations." Cancer Epidemiol Biomarkers Prev **10**(12): 1239-48.

Geddes, D. M. (1979). "The natural history of lung cancer: a review based on rates of tumour growth." Br J Dis Chest **73**(1): 1-17.

Gedik, C. M., S. P. Boyle, S. G. Wood, N. J. Vaughan and A. R. Collins (2002). "Oxidative stress in humans: validation of biomarkers of DNA damage." Carcinogenesis 23(9): 1441-6.

Gedik, C. M. and A. Collins (2005). "Establishing the background level of base oxidation in human lymphocyte DNA: results of an interlaboratory validation study." Faseb J 19(1): 82-4.

Georgiadis, P., E. Samoli, S. Kaila, K. Katsouyanni and S. A. Kyrtopoulos (2000). "Ubiquitous presence of O6-methylguanine in human peripheral and cord blood DNA." Cancer Epidemiol Biomarkers Prev 9(3): 299-305.

Gerson, S. L., J. E. Trey, K. Miller and N. A. Berger (1986). "Comparison of O6-alkylguanine-DNA alkyltransferase activity based on cellular DNA content in human, rat and mouse tissues." Carcinogenesis 7(5): 745-9.

Ginsberg, R. J. and L. V. Rubinstein (1995). "Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. Lung Cancer Study Group." Ann Thorac Surg 60(3): 615-22; discussion 622-3.

Godtfredsen, N. S., E. Prescott and M. Osler (2005). "Effect of smoking reduction on lung cancer risk." Jama 294(12): 1505-10.

Goldman, R. and P. G. Shields (2003). "Food mutagens." J Nutr 133 Suppl 3: 965S-973S.

Grafstrom, R. C., A. E. Pegg, B. F. Trump and C. C. Harris (1984). "O6-alkylguanine-DNA alkyltransferase activity in normal human tissues and cells." Cancer Res 44(7): 2855-7.

Greenberg, A. K., H. Yee and W. N. Rom (2002). "Preneoplastic lesions of the lung." Respir Res 3(1): 20.

Grombacher, T. and B. Kaina (1995). "Constitutive expression and inducibility of O6-methylguanine-DNA methyltransferase and N-methylpurine-DNA glycosylase in rat liver cells exhibiting different status of differentiation." Biochim Biophys Acta 1270(1): 63-72.

Grosche, B., M. Kreuzer, M. Kreisheimer, M. Schnelzer and A. Tschense (2006). "Lung cancer risk among German male uranium miners: a cohort study, 1946-1998." Br J Cancer 95(9): 1280-1287.

Group, E. L. C. (1987). "Patients presenting with lung cancer in south east Scotland. Edinburgh Lung Cancer Group." Thorax 42(11): 853-7.

Hackshaw, A. K., M. R. Law and N. J. Wald (1997). "The accumulated evidence on lung cancer and environmental tobacco smoke." Bmj 315(7114): 980-8.

Hall, J., H. Bresil, F. Donato, C. P. Wild, N. A. Loktionova, O. I. Kazanova, I. P. Komyakov, V. G. Lemekhov, A. J. Likhachev and R. Montesano (1993). "Alkylation and oxidative-DNA damage repair activity in blood leukocytes of smokers and non-smokers." Int J Cancer 54(5): 728-33.

Hall, J., H. Bresil, M. Serres, G. Martel-Planche, C. P. Wild and R. Montesano (1990). "Modulation of O6-methylguanine-DNA methyltransferase in rat and hamster liver after treatment with dimethylnitrosamine." Cancer Res 50(17): 5426-30.

Han, J., S. E. Hankinson and I. De Vivo (2006). "Polymorphisms in O6-methylguanine DNA methyltransferase and endometrial cancer risk." Carcinogenesis.

Han, J., G. J. Tranah, S. E. Hankinson, L. D. Samson and D. J. Hunter (2006). "Polymorphisms in O6-methylguanine DNA methyltransferase and breast cancer risk." Pharmacogenet Genomics 16(7): 469-474.

Hanahan, D. and R. A. Weinberg (2000). "The hallmarks of cancer." Cell 100(1): 57-70.

Hang, B. and B. Singer (2003). "Protein-protein interactions involving DNA glycosylases." Chem Res Toxicol 16(10): 1181-95.

Hang, B., B. Singer, G. P. Margison and R. H. Elder (1997). "Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1,N6-ethenoadenine and hypoxanthine but not of 3,N4-ethenocytosine or 8-oxoguanine." Proc Natl Acad Sci U S A 94(24): 12869-74.

Hardie, L. J., J. A. Briggs, L. A. Davidson, J. M. Allan, R. F. King, G. I. Williams and C. P. Wild (2000). "The effect of hOGG1 and glutathione peroxidase I genotypes and 3p chromosomal loss on 8-hydroxydeoxyguanosine levels in lung cancer." Carcinogenesis 21(2): 167-72.

Harrison, K. L., P. A. Crosbie, R. M. Agius, P. V. Barber, M. Carus, G. P. Margison and A. C. Povey (2006). "No association between N7-methyldeoxyguanosine and 8-oxodeoxyguanosine levels in human lymphocyte DNA." Mutat Res.

Harrison, K. L., M. Wood, N. P. Lees, C. N. Hall, G. P. Margison and A. C. Povey (2001). "Development and application of a sensitive and rapid immunoassay for the quantitation of N7-methyldeoxyguanosine in DNA samples." Chem Res Toxicol 14(3): 295-301.

Hayes, J. D., J. U. Flanagan and I. R. Jowsey (2005). "Glutathione transferases." Annu Rev Pharmacol Toxicol 45: 51-88.

Hayes, J. D. and R. C. Strange (2000). "Glutathione S-transferase polymorphisms and their biological consequences." Pharmacology 61(3): 154-66.

Hecht, S. S. (1996). "Chemoprevention of lung cancer by isothiocyanates." Adv Exp Med Biol 401: 1-11.

Hecht, S. S. (1998). "Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines." Chem Res Toxicol 11(6): 559-603.

Hecht, S. S. (1999). "DNA adduct formation from tobacco-specific N-nitrosamines." Mutat Res 424(1-2): 127-42.

Hecht, S. S. (1999). "Tobacco smoke carcinogens and lung cancer." J Natl Cancer Inst 91(14): 1194-210.

Hecht, S. S. (2002). "Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention." Lancet Oncol 3(8): 461-9.

Heighway, J., G. P. Margison and M. F. Santibanez-Koref (2003). "The alleles of the DNA repair gene O6-alkylguanine-DNA alkyltransferase are expressed at different levels in normal human lung tissue." Carcinogenesis 24(10): 1691-4.

Henley, S. J., M. J. Thun, A. Chao and E. E. Calle (2004). "Association between exclusive pipe smoking and mortality from cancer and other diseases." J Natl Cancer Inst 96(11): 853-61.

Hennekens, C. H., J. E. Buring, J. E. Manson, M. Stampfer, B. Rosner, N. R. Cook, C. Belanger, F. LaMotte, J. M. Gaziano, P. M. Ridker, W. Willett and R. Peto (1996). "Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease." <u>N Engl J Med</u> **334**(18): 1145-9.

Henschke, C. I., D. I. McCauley, D. F. Yankelevitz, D. P. Naidich, G. McGuinness, O. S. Miettinen, D. M. Libby, M. W. Pasmantier, J. Koizumi, N. K. Altorki and J. P. Smith (1999). "Early Lung Cancer Action Project: overall design and findings from baseline screening." <u>Lancet</u> **354**(9173): 99-105.

Henschke, C. I. and O. S. Miettinen (2004). "Women's susceptibility to tobacco carcinogens." <u>Lung Cancer</u> **43**(1): 1-5.

Hershberger, P. A., A. C. Vasquez, B. Kanterewicz, S. Land, J. M. Siegfried and M. Nichols (2005). "Regulation of endogenous gene expression in human non-small cell lung cancer cells by estrogen receptor ligands." <u>Cancer Res</u> **65**(4): 1598-605.

Hill, C. E., A. A. Affatato, K. J. Wolfe, M. S. Lopez, C. K. Hallberg, D. Canistro and S. Z. Abdel-Rahman (2005). "Gender differences in genetic damage induced by the tobacco-specific nitrosamine NNK and the influence of the Thr241Met polymorphism in the XRCC3 gene." <u>Environ Mol Mutagen</u> **46**(1): 22-9.

Hill, J. W., T. K. Hazra, T. Izumi and S. Mitra (2001). "Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair." <u>Nucleic Acids Res</u> **29**(2): 430-8.

Hiltonen, T., A. K. Clarke, J. Karlsson and G. Samuelsson (1996). "A cDNA coding for glutathione S-transferase from the unicellular green algae Coccomyxa sp." <u>Gene</u> **176**(1-2): 263-4.

Hoeijmakers, J. H. (2001). "Genome maintenance mechanisms for preventing cancer." <u>Nature</u> **411**(6835): 366-74.

Hofer, T. and L. Moller (2002). "Optimization of the workup procedure for the analysis of 8-oxo-7,8-dihydro-2'-deoxyguanosine with electrochemical detection." <u>Chem Res Toxicol</u> **15**(3): 426-32.

Hoffmann, D., M. V. Djordjevic and I. Hoffmann (1997). "The changing cigarette." <u>Prev Med</u> **26**(4): 427-34.

Holland, J. M., A. Arsanjani, B. J. Liem, S. C. Hoffelt, J. I. Cohen and K. R. Stevens, Jr. (2002). "Second malignancies in early stage laryngeal carcinoma patients treated with radiotherapy." <u>J Laryngol Otol</u> **116**(3): 190-3.

Hotchkiss, J. H. (1987). "A review of current literature on N-nitroso compounds in foods." <u>Adv Food Res</u> **31**: 53-115.

Houlston, R. S. (1999). "Glutathione S-transferase M1 status and lung cancer risk: a meta-analysis." <u>Cancer Epidemiol Biomarkers Prev</u> **8**(8): 675-82.

Houlston, R. S. (2000). "CYP1A1 polymorphisms and lung cancer risk: a meta-analysis." <u>Pharmacogenetics</u> **10**(2): 105-14.

Hubbard, R., A. Venn, S. Lewis and J. Britton (2000). "Lung cancer and cryptogenic fibrosing alveolitis. A population-based cohort study." Am J Respir Crit Care Med **161**(1): 5-8.

Hung, R. J., P. Brennan, F. Canzian, N. Szeszenia-Dabrowska, D. Zaridze, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, L. Foretova, V. Janout, V. Bencko, A. Chabrier, S. Borel, J. Hall and P. Boffetta (2005). "Large-scale investigation of base excision repair genetic polymorphisms and lung cancer risk in a multicenter study." J Natl Cancer Inst **97**(8): 567-76.

Ingelman-Sundberg, M., I. Johansson, I. Persson, Q. Y. Yue, M. L. Dahl, L. Bertilsson and F. Sjoqvist (1992). "Genetic polymorphism of cytochromes P450: interethnic differences and relationship to incidence of lung cancer." Pharmacogenetics **2**(6): 264-71.

Inoue, M., T. Osaki, M. Noguchi, S. Hirohashi, K. Yasumoto and H. Kasai (1998). "Lung cancer patients have increased 8-hydroxydeoxyguanosine levels in peripheral lung tissue DNA." Jpn J Cancer Res **89**(7): 691-5.

Iribarren, C., I. S. Tekawa, S. Sidney and G. D. Friedman (1999). "Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men." N Engl J Med **340**(23): 1773-80.

Ishibashi, T., Y. Nakabeppu, H. Kawate, K. Sakumi, H. Hayakawa and M. Sekiguchi (1994). "Intracellular localization and function of DNA repair methyltransferase in human cells." Mutat Res **315**(3): 199-212.

Ivaschenko, T. E., O. G. Sideleva and V. S. Baranov (2002). "Glutathione- S-transferase micro and theta gene polymorphisms as new risk factors of atopic bronchial asthma." J Mol Med **80**(1): 39-43.

Iwakuma, T., K. Sakumi, Y. Nakatsuru, H. Kawate, H. Igarashi, A. Shiraishi, T. Tsuzuki, T. Ishikawa and M. Sekiguchi (1997). "High incidence of nitrosamine-induced tumorigenesis in mice lacking DNA repair methyltransferase." Carcinogenesis **18**(8): 1631-5.

Jacobson, D. R. (1999). ras Mutations in Lung Cancer. Lung Tumours Fundamental Biology and Clinical Management. E. Brambilla and C. Brambilla, Marcel Dekker. **124:** 139-152.

Janne, P. A., B. Freidlin, S. Saxman, D. H. Johnson, R. B. Livingston, F. A. Shepherd and B. E. Johnson (2002). "Twenty-five years of clinical research for patients with limited-stage small cell lung carcinoma in North America." Cancer **95**(7): 1528-38.

Janne, P. A., S. Gurubhagavatula, B. Y. Yeap, J. Lucca, P. Ostler, A. T. Skarin, P. Fidias, T. J. Lynch and B. E. Johnson (2004). "Outcomes of patients with advanced non-small cell lung cancer treated with gefitinib (ZD1839, "Iressa") on an expanded access study." Lung Cancer **44**(2): 221-30.

Janssen, K., U. Eichhorn-Grombacher, K. Schlink, S. Nitzsche, F. Oesch and B. Kaina (2001). "Long-time expression of DNA repair enzymes MGMT and APE in human peripheral blood mononuclear cells." Arch Toxicol **75**(5): 306-12.

Janssen-Heijnen, M. L. and J. W. Coebergh (2003). "The changing epidemiology of lung cancer in Europe." Lung Cancer **41**(3): 245-58.

Jaruga, P., T. H. Zastawny, J. Skokowski, M. Dizdaroglu and R. Olinski (1994). "Oxidative DNA base damage and antioxidant enzyme activities in human lung cancer." FEBS Lett **341**(1): 59-64.

Kamataki, T., M. Fujieda, K. Kiyotani, S. Iwano and H. Kunitoh (2005). "Genetic polymorphism of CYP2A6 as one of the potential determinants of tobacco-related cancer risk." Biochem Biophys Res Commun 338(1): 306-10.

Kamataki, T., K. Nunoya, Y. Sakai, H. Kushida and K. Fujita (1999). "Genetic polymorphism of CYP2A6 in relation to cancer." Mutat Res 428(1-2): 125-30.

Karran, P. a. L., T (1985). Cellular defence mechanisms against alkylating agents. Cancer Surveys. 4: 583-599.

Kasai, H. (1997). "Analysis of a form of oxidative DNA damage, 8-hydroxy-2'-deoxyguanosine, as a marker of cellular oxidative stress during carcinogenesis." Mutat Res 387(3): 147-63.

Kato, S., E. D. Bowman, A. M. Harrington, B. Blomeke and P. G. Shields (1995). "Human lung carcinogen-DNA adduct levels mediated by genetic polymorphisms in vivo." J Natl Cancer Inst 87(12): 902-7.

Kato, S., S. Petruzzelli, E. D. Bowman, K. W. Turteltaub, B. Blomeke, A. Weston and P. G. Shields (1993). "7-Alkyldeoxyguanosine adduct detection by two-step HPLC and the 32P-postlabeling assay." Carcinogenesis 14(4): 545-50.

Kaur, T. B., J. M. Travaline, J. P. Gaughan, J. P. Richie, Jr., S. D. Stellman and P. Lazarus (2000). "Role of polymorphisms in codons 143 and 160 of the O6-alkylguanine DNA alkyltransferase gene in lung cancer risk." Cancer Epidemiol Biomarkers Prev 9(3): 339-42.

Khuder, S. A. and A. B. Mutgi (2001). "Effect of smoking cessation on major histologic types of lung cancer." Chest 120(5): 1577-83.

Kim, N. K., J. Y. Ahn, J. Song, J. K. Kim, J. H. Han, H. J. An, H. M. Chung, J. Y. Joo, J. U. Choi, K. S. Lee, R. Roy and D. Oh (2003). "Expression of the DNA repair enzyme, N-methylpurine-DNA glycosylase (MPG) in astrocytic tumors." Anticancer Res 23(2B): 1417-23.

Kim, N. K., H. J. An, H. J. Kim, T. J. Sohn, R. Roy, D. Oh, J. Y. Ahn, T. S. Hwang and K. Y. Cha (2002). "Altered expression of the DNA repair protein, N-methylpurine-DNA glycosylase (MPG) in human gonads." Anticancer Res 22(2A): 793-8.

Kim, N. K., S. H. Lee, K. Y. Cha and J. S. Seo (1998). "Tissue-specific expression and activation of N-methylpurine-DNA glycosylase in thymic carcinomas of transgenic mice expressing the SV40 large T-antigen gene." Mol Cells 8(4): 383-7.

Kishi, K., J. W. Gurney, D. R. Schroeder, P. D. Scanlon, S. J. Swensen and J. R. Jett (2002). "The correlation of emphysema or airway obstruction with the risk of lung cancer: a matched case-controlled study." Eur Respir J 19(6): 1093-8.

Kiyohara, C., A. Otsu, T. Shirakawa, S. Fukuda and J. M. Hopkin (2002). "Genetic polymorphisms and lung cancer susceptibility: a review." Lung Cancer 37(3): 241-56.

Kiyohara, C., K. I. Yamamura, Y. Nakanishi, K. Takayama and N. Hara (2000). "Polymorphism in GSTM1, GSTT1, and GSTP1 and Susceptibility to Lung Cancer in a Japanese Population." Asian Pac J Cancer Prev 1(4): 293-298.

Kiyosawa, H., M. Suko, H. Okudaira, K. Murata, T. Miyamoto, M. H. Chung, H. Kasai and S. Nishimura (1990). "Cigarette smoking induces formation of 8-hydroxydeoxyguanosine, one of

the oxidative DNA damages in human peripheral leukocytes." Free Radic Res Commun 11(1-3): 23-7.

Klaunig, J. E. and L. M. Kamendulis (2004). "The role of oxidative stress in carcinogenesis." Annu Rev Pharmacol Toxicol 44: 239-67.

Klungland, A., I. Rosewell, S. Hollenbach, E. Larsen, G. Daly, B. Epe, E. Seeberg, T. Lindahl and D. E. Barnes (1999). "Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage." Proc Natl Acad Sci U S A 96(23): 13300-5.

Kohno, T., H. Kunitoh, K. Toyama, S. Yamamoto, A. Kuchiba, D. Saito, N. Yanagitani, S. Ishihara, R. Saito and J. Yokota (2006). "Association of the OGG1-Ser326Cys polymorphism with lung adenocarcinoma risk." Cancer Sci 97(8): 724-8.

Kreklau, E. L., M. Limp-Foster, N. Liu, Y. Xu, M. R. Kelley and L. C. Erickson (2001). "A novel fluorometric oligonucleotide assay to measure O(6)-methylguanine DNA methyltransferase, methylpurine DNA glycosylase, 8-oxoguanine DNA glycosylase and abasic endonuclease activities: DNA repair status in human breast carcinoma cells overexpressing methylpurine DNA glycosylase." Nucleic Acids Res 29(12): 2558-66.

Kreuzer, M., P. Boffetta, E. Whitley, W. Ahrens, V. Gaborieau, J. Heinrich, K. H. Jockel, L. Kreienbrock, S. Mallone, F. Merletti, F. Roesch, P. Zambon and L. Simonato (2000). "Gender differences in lung cancer risk by smoking: a multicentre case-control study in Germany and Italy." Br J Cancer 82(1): 227-33.

Kreuzer, M., M. Gerken, J. Heinrich, L. Kreienbrock and H. E. Wichmann (2003). "Hormonal factors and risk of lung cancer among women?" Int J Epidemiol 32(2): 263-71.

Krzesniak, M., D. Butkiewicz, A. Samojedny, M. Chorazy and M. Rusin (2004). "Polymorphisms in TDG and MGMT genes - epidemiological and functional study in lung cancer patients from Poland." Ann Hum Genet 68(Pt 4): 300-12.

Kubik, A. and J. Polak (1986). "Lung cancer detection. Results of a randomized prospective study in Czechoslovakia." Cancer 57(12): 2427-37.

Kumar, R. and K. Hemminki (1996). "Separation of 7-methyl- and 7-(2-hydroxyethyl)-guanine adducts in human DNA samples using a combination of TLC and HPLC." Carcinogenesis 17(3): 485-92.

Kyrtopoulos, S. A. (1998). "DNA adducts in humans after exposure to methylating agents." Mutat Res 405(2): 135-43.

Landi, S. (2000). "Mammalian class theta GST and differential susceptibility to carcinogens: a review." Mutat Res 463(3): 247-83.

Larson, K., J. Sahm, R. Shenkar and B. Strauss (1985). "Methylation-induced blocks to in vitro DNA replication." Mutat Res 150(1-2): 77-84.

Le Marchand, L., T. Donlon, A. Lum-Jones, A. Seifried and L. R. Wilkens (2002). "Association of the hOGG1 Ser326Cys polymorphism with lung cancer risk." Cancer Epidemiol Biomarkers Prev 11(4): 409-12.

Lee, S. M., J. A. Rafferty, R. H. Elder, C. Y. Fan, M. Bromley, M. Harris, N. Thatcher, P. M. Potter, H. J. Altermatt, T. Perinat-Frey and et al. (1992). "Immunohistological examination of

the inter- and intracellular distribution of O6-alkylguanine DNA-alkyltransferase in human liver and melanoma." Br J Cancer **66**(2): 355-60.

Lengauer, C., K. W. Kinzler and B. Vogelstein (1998). "Genetic instabilities in human cancers." Nature **396**(6712): 643-9.

Lenton, K. J., H. Therriault, T. Fulop, H. Payette and J. R. Wagner (1999). "Glutathione and ascorbate are negatively correlated with oxidative DNA damage in human lymphocytes." Carcinogenesis **20**(4): 607-13.

Lewis, S. J., N. M. Cherry, L. N. R. Mc, P. V. Barber, K. Wilde and A. C. Povey (2003). "Cotinine levels and self-reported smoking status in patients attending a bronchoscopy clinic." Biomarkers **8**(3-4): 218-28.

Lewis, S. J., N. M. Cherry, R. M. Niven, P. V. Barber and A. C. Povey (2002). "GSTM1, GSTT1 and GSTP1 polymorphisms and lung cancer risk." Cancer Lett **180**(2): 165-71.

Lewis, S. J., N. M. Cherry, R. M. Niven, P. V. Barber and A. C. Povey (2004). "Associations between smoking, GST genotypes and N7-methylguanine levels in DNA extracted from bronchial lavage cells." Mutat Res **559**(1-2): 11-8.

Li, D., P. F. Firozi, L. E. Wang, C. H. Bosken, M. R. Spitz, W. K. Hong and Q. Wei (2001). "Sensitivity to DNA damage induced by benzo(a)pyrene diol epoxide and risk of lung cancer: a case-control analysis." Cancer Res **61**(4): 1445-50.

Likhite, V. S., E. I. Cass, S. D. Anderson, J. R. Yates and A. M. Nardulli (2004). "Interaction of estrogen receptor alpha with 3-methyladenine DNA glycosylase modulates transcription and DNA repair." J Biol Chem **279**(16): 16875-82.

Lim, A. and B. F. Li (1996). "The nuclear targeting and nuclear retention properties of a human DNA repair protein O6-methylguanine-DNA methyltransferase are both required for its nuclear localization: the possible implications." Embo J **15**(15): 4050-60.

Lindahl, T. and D. E. Barnes (2000). "Repair of endogenous DNA damage." Cold Spring Harb Symp Quant Biol **65**: 127-33.

Liu, L., X. Qin and S. L. Gerson (1999). "Reduced lung tumorigenesis in human methylguanine DNA--methyltransferase transgenic mice achieved by expression of transgene within the target cell." Carcinogenesis **20**(2): 279-84.

Lodovici, M., C. Casalini, R. Cariaggi, L. Michelucci and P. Dolara (2000). "Levels of 8-hydroxydeoxyguanosine as a marker of DNA damage in human leukocytes." Free Radic Biol Med **28**(1): 13-7.

Loft, S., P. Svoboda, H. Kasai, A. Tjonneland, U. Vogel, P. Moller, K. Overvad and O. Raaschou-Nielsen (2005). "Prospective study of 8-oxo-7,8-dihydro-2'-deoxyguanosine excretion and the risk of lung cancer." Carcinogenesis.

London, S. J., J. M. Yuan, F. L. Chung, Y. T. Gao, G. A. Coetzee, R. K. Ross and M. C. Yu (2000). "Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms, and lung-cancer risk: a prospective study of men in Shanghai, China." Lancet **356**(9231): 724-9.

Loriot, M. A., S. Rebuissou, M. Oscarson, S. Cenee, M. Miyamoto, N. Ariyoshi, T. Kamataki, D. Hemon, P. Beaune and I. Stucker (2001). "Genetic polymorphisms of cytochrome P450 2A6 in a case-control study on lung cancer in a French population." Pharmacogenetics 11(1): 39-44.

Lubin, J. H. and J. D. Boice, Jr. (1997). "Lung cancer risk from residential radon: meta-analysis of eight epidemiologic studies." J Natl Cancer Inst 89(1): 49-57.

Maher, R. L., A. C. Vallur, J. A. Feller and L. B. Bloom (2007). "Slow base excision by human alkyladenine DNA glycosylase limits the rate of formation of AP sites and AP endonuclease 1 does not stimulate base excision." DNA Repair (Amst) 6(1): 71-81.

Marchenay, C., E. Cellarier, F. Levi, C. Rolhion, F. Kwiatkowski, B. Claustrat, J. C. Madelmont and P. Chollet (2001). "Circadian variation in O6-alkylguanine-DNA alkyltransferase activity in circulating blood mononuclear cells of healthy human subjects." Int J Cancer 91(1): 60-6.

Margison, G. P., J. Heighway, S. Pearson, G. McGown, M. R. Thorncroft, A. J. Watson, K. L. Harrison, S. J. Lewis, K. Rohde, P. V. Barber, P. O'Donnell, A. C. Povey and M. F. Santibanez-Koref (2005). "Quantitative trait locus analysis reveals two intragenic sites that influence O6-alkylguanine-DNA alkyltransferase activity in peripheral blood mononuclear cells." Carcinogenesis 26(8): 1473-80.

Margison, G. P., A. C. Povey, B. Kaina and M. F. Santibanez Koref (2003). "Variability and regulation of O6-alkylguanine-DNA alkyltransferase." Carcinogenesis 24(4): 625-35.

Margison, G. P., M. F. Santibanez Koref and A. C. Povey (2002). "Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine." Mutagenesis 17(6): 483-7.

Marnett, L. J. and P. C. Burcham (1993). "Endogenous DNA adducts: potential and paradox." Chem Res Toxicol 6(6): 771-85.

Maser, E. (2004). "Significance of reductases in the detoxification of the tobacco-specific carcinogen NNK." Trends Pharmacol Sci 25(5): 235-7.

Masuda, Y., R. A. Bennett and B. Demple (1998). "Rapid dissociation of human apurinic endonuclease (Ape1) from incised DNA induced by magnesium." J Biol Chem 273(46): 30360-5.

Matakidou, A., T. Eisen and R. S. Houlston (2005). "Systematic review of the relationship between family history and lung cancer risk." Br J Cancer 93(7): 825-33.

Mattern, J., R. Koomagi and M. Volm (1998). "Smoking-related increase of O6-methylguanine-DNA methyltransferase expression in human lung carcinomas." Carcinogenesis 19(7): 1247-50.

McGarry, R. C., G. Song, P. des Rosiers and R. Timmerman (2002). "Observation-only management of early stage, medically inoperable lung cancer: poor outcome." Chest 121(4): 1155-8.

McLellan, R. A., M. Oscarson, A. K. Alexandrie, J. Seidegard, D. A. Evans, A. Rannug and M. Ingelman-Sundberg (1997). "Characterization of a human glutathione S-transferase mu cluster containing a duplicated GSTM1 gene that causes ultrarapid enzyme activity." Mol Pharmacol 52(6): 958-65.

Melamed, M. R., B. J. Flehinger, M. B. Zaman, R. T. Heelan, W. A. Perchick and N. Martini (1984). "Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York." Chest 86(1): 44-53.

Mijal, R. S., S. Kanugula, C. C. Vu, Q. Fang, A. E. Pegg and L. A. Peterson (2006). "DNA sequence context affects repair of the tobacco-specific adduct O(6)-[4-Oxo-4-(3-pyridyl)butyl]guanine by human O(6)-alkylguanine-DNA alkyltransferases." Cancer Res 66(9): 4968-74.

Miller, A. B., H. P. Altenburg, B. Bueno-de-Mesquita, H. C. Boshuizen, A. Agudo, F. Berrino, I. T. Gram, L. Janson, J. Linseisen, K. Overvad, T. Rasmuson, P. Vineis, A. Lukanova, N. Allen, P. Amiano, A. Barricarte, G. Berglund, H. Boeing, F. Clavel-Chapelon, N. E. Day, G. Hallmans, E. Lund, C. Martinez, C. Navarro, D. Palli, S. Panico, P. H. Peeters, J. R. Quiros, A. Tjonneland, R. Tumino, A. Trichopoulou, D. Trichopoulos, N. Slimani and E. Riboli (2004). "Fruits and vegetables and lung cancer: Findings from the European Prospective Investigation into Cancer and Nutrition." Int J Cancer 108(2): 269-76.

Mollerup, S., G. Berge, R. Baera, V. Skaug, A. Hewer, D. H. Phillips, L. Stangeland and A. Haugen (2006). "Sex differences in risk of lung cancer: Expression of genes in the PAH bioactivation pathway in relation to smoking and bulky DNA adducts." Int J Cancer 119(4): 741-4.

Mollerup, S., D. Ryberg, A. Hewer, D. H. Phillips and A. Haugen (1999). "Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients." Cancer Res 59(14): 3317-20.

Morland, I., L. Luna, E. Gustad, E. Seeberg and M. Bjoras (2005). "Product inhibition and magnesium modulate the dual reaction mode of hOgg1." DNA Repair (Amst) 4(3): 381-7.

Mountain, C. F. (1997). "Revisions in the International System for Staging Lung Cancer." Chest 111(6): 1710-7.

Muhm, J. R., W. E. Miller, R. S. Fontana, D. R. Sanderson and M. A. Uhlenhopp (1983). "Lung cancer detected during a screening program using four-month chest radiographs." Radiology 148(3): 609-15.

Mustonen, R. and K. Hemminki (1992). "7-Methylguanine levels in DNA of smokers' and non-smokers' total white blood cells, granulocytes and lymphocytes." Carcinogenesis 13(11): 1951-5.

Mustonen, R., B. Schoket and K. Hemminki (1993). "Smoking-related DNA adducts: 32P-postlabeling analysis of 7-methylguanine in human bronchial and lymphocyte DNA." Carcinogenesis 14(1): 151-4.

Nakajima, T., E. Elovaara, S. Anttila, A. Hirvonen, A. M. Camus, J. D. Hayes, B. Ketterer and H. Vainio (1995). "Expression and polymorphism of glutathione S-transferase in human lungs: risk factors in smoking-related lung cancer." Carcinogenesis 16(4): 707-11.

Nakamura, J. and J. A. Swenberg (1999). "Endogenous apurinic/apyrimidinic sites in genomic DNA of mammalian tissues." Cancer Res 59(11): 2522-6.

Nakatsuru, Y., E. Tsuchiya, K. Nakagawa, N. Nemoto, H. Oda and T. Ishikawa (1994). "O6-methylguanine-DNA methyltransferase activity in the human lung persists with advancing age." Gerontology 40 Suppl 2: 3-9.

279

Nations, F. a. A. O. U. (2003). Projections of tobacco production, consumption and trade to year 2010. Rome, United Nations: 1-101.

Nazar-Stewart, V., T. L. Vaughan, P. Stapleton, J. Van Loo, B. Nicol-Blades and D. L. Eaton (2003). "A population-based study of glutathione S-transferase M1, T1 and P1 genotypes and risk for lung cancer." Lung Cancer 40(3): 247-58.

Nelson, H. H., D. C. Christiani, E. J. Mark, J. K. Wiencke, J. C. Wain and K. T. Kelsey (1999). "Implications and prognostic value of K-ras mutation for early-stage lung cancer in women." J Natl Cancer Inst 91(23): 2032-8.

Nelson, H. H., J. K. Wiencke, D. C. Christiani, T. J. Cheng, Z. F. Zuo, B. S. Schwartz, B. K. Lee, M. R. Spitz, M. Wang, X. Xu and et al. (1995). "Ethnic differences in the prevalence of the homozygous deleted genotype of glutathione S-transferase theta." Carcinogenesis 16(5): 1243-5.

O'Brien, P. J. and T. Ellenberger (2003). "Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines." Biochemistry 42(42): 12418-29.

O'Brien, P. J. and T. Ellenberger (2004). "The Escherichia coli 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site." J Biol Chem 279(26): 26876-84.

O'Donnell, P. N., P. V. Barber, G. P. Margison and A. C. Povey (1999). "Association between O6-alkylguanine-DNA-alkyltransferase activity in peripheral blood lymphocytes and bronchial epithelial cells." Cancer Epidemiol Biomarkers Prev 8(7): 641-5.

Oken, M. M., P. M. Marcus, P. Hu, T. M. Beck, W. Hocking, P. A. Kvale, J. Cordes, T. L. Riley, S. D. Winslow, S. Peace, D. L. Levin, P. C. Prorok and J. K. Gohagan (2005). "Baseline chest radiograph for lung cancer detection in the randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial." J Natl Cancer Inst 97(24): 1832-9.

Olinski, R., D. Gackowski, R. Rozalski, M. Foksinski and K. Bialkowski (2003). "Oxidative DNA damage in cancer patients: a cause or a consequence of the disease development?" Mutat Res 531(1-2): 177-90.

Omenn, G. S., G. E. Goodman, M. D. Thornquist, J. Balmes, M. R. Cullen, A. Glass, J. P. Keogh, F. L. Meyskens, B. Valanis, J. H. Williams, S. Barnhart and S. Hammar (1996). "Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease." N Engl J Med 334(18): 1150-5.

Omoto, Y., Y. Kobayashi, K. Nishida, E. Tsuchiya, H. Eguchi, K. Nakagawa, Y. Ishikawa, T. Yamori, H. Iwase, Y. Fujii, M. Warner, J. A. Gustafsson and S. I. Hayashi (2001). "Expression, function, and clinical implications of the estrogen receptor beta in human lung cancers." Biochem Biophys Res Commun 285(2): 340-7.

Ooi, W. L., R. C. Elston, V. W. Chen, J. E. Bailey-Wilson and H. Rothschild (1986). "Increased familial risk for lung cancer." J Natl Cancer Inst 76(2): 217-22.

Osterod, M., S. Hollenbach, J. G. Hengstler, D. E. Barnes, T. Lindahl and B. Epe (2001). "Age-related and tissue-specific accumulation of oxidative DNA base damage in 7,8-dihydro-8-oxoguanine-DNA glycosylase (Ogg1) deficient mice." Carcinogenesis 22(9): 1459-63.

Paez, J. G., P. A. Janne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M. J. Eck, W. R. Sellers, B. E. Johnson

and M. Meyerson (2004). "EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy." Science **304**(5676): 1497-500.

Paik, J., T. Duncan, T. Lindahl and B. Sedgwick (2005). "Sensitization of human carcinoma cells to alkylating agents by small interfering RNA suppression of 3-alkyladenine-DNA glycosylase." Cancer Res **65**(22): 10472-7.

Pandya, G. A. and M. Moriya (1996). "1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells." Biochemistry **35**(35): 11487-92.

Pant, P. V., H. Tao, E. J. Beilharz, D. G. Ballinger, D. R. Cox and K. A. Frazer (2006). "Analysis of allelic differential expression in human white blood cells." Genome Res **16**(3): 331-9.

Park, J., L. Chen, M. S. Tockman, A. Elahi and P. Lazarus (2004). "The human 8-oxoguanine DNA N-glycosylase 1 (hOGG1) DNA repair enzyme and its association with lung cancer risk." Pharmacogenetics **14**(2): 103-9.

Parkin, D. M., F. Bray, J. Ferlay and P. Pisani (2005). "Global cancer statistics, 2002." CA Cancer J Clin **55**(2): 74-108.

Parkin, D. M. and S. M. Moss (2000). "Lung cancer screening: improved survival but no reduction in deaths--the role of "overdiagnosis"." Cancer **89**(11 Suppl): 2369-76.

Parl, F. F. (2005). "Glutathione S-transferase genotypes and cancer risk." Cancer Lett **221**(2): 123-9.

Parsons, J. L. and R. H. Elder (2003). "DNA N-glycosylase deficient mice: a tale of redundancy." Mutat Res **531**(1-2): 165-75.

Pastinen, T. and T. J. Hudson (2004). "Cis-acting regulatory variation in the human genome." Science **306**(5696): 647-50.

Patz, E. F., Jr., P. C. Goodman and G. Bepler (2000). "Screening for lung cancer." N Engl J Med **343**(22): 1627-33.

Paz-Elizur, T., M. Krupsky, S. Blumenstein, D. Elinger, E. Schechtman and Z. Livneh (2003). "DNA repair activity for oxidative damage and risk of lung cancer." J Natl Cancer Inst **95**(17): 1312-9.

Pegg, A. E. (2000). "Repair of O(6)-alkylguanine by alkyltransferases." Mutat Res **462**(2-3): 83-100.

Peterson, L. A. and S. S. Hecht (1991). "O6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung." Cancer Res **51**(20): 5557-64.

Peterson, L. A., X. K. Liu and S. S. Hecht (1993). "Pyridyloxobutyl DNA adducts inhibit the repair of O6-methylguanine." Cancer Res **53**(12): 2780-5.

Peto, R., S. Darby, H. Deo, P. Silcocks, E. Whitley and R. Doll (2000). "Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies." Bmj **321**(7257): 323-9.

Pieper, R. O. (1997). "Understanding and manipulating O6-methylguanine-DNA methyltransferase expression." Pharmacol Ther 74(3): 285-97.

Plosky, B., L. Samson, B. P. Engelward, B. Gold, B. Schlaen, T. Millas, M. Magnotti, J. Schor and D. A. Scicchitano (2002). "Base excision repair and nucleotide excision repair contribute to the removal of N-methylpurines from active genes." DNA Repair (Amst) 1(8): 683-96.

Plummer, E. R. (2006). "Inhibition of poly(ADP-ribose) polymerase in cancer." Curr Opin Pharmacol 6(4): 364-8.

Poulsen, H. E. (2005). "Oxidative DNA modifications." Exp Toxicol Pathol 57 Suppl 1: 161-9.

Pourcelot, S., H. Faure, F. Firoozi, V. Ducros, M. Tripier, J. Hee, J. Cadet and A. Favier (1999). "Urinary 8-oxo-7,8-dihydro-2'-deoxyguanosine and 5-(hydroxymethyl) uracil in smokers." Free Radic Res 30(3): 173-80.

Povey, A. C., P. O'Donnell, P. Barber, M. Watson, G. P. Margison and M. F. Santibanez Koref (2006). "Smoking is associated with a decrease of O6-alkylguanine-DNA alkyltransferase activity in bronchial epithelial cells." Int J Cancer 119(2): 463-6.

Pujol, J. L., L. Carestia and J. P. Daures (2000). "Is there a case for cisplatin in the treatment of small-cell lung cancer? A meta-analysis of randomized trials of a cisplatin-containing regimen versus a regimen without this alkylating agent." Br J Cancer 83(1): 8-15.

Pulera, N., S. Petruzzelli, A. Celi, R. Puntoni, E. Fornai, U. Sawe, P. Paoletti and C. Giuntini (1997). "Presence and persistence of serum anti-benzo[a]pyrene diolepoxide-DNA adduct antibodies in smokers: effects of smoking reduction and cessation." Int J Cancer 70(2): 145-9.

Pulling, L. C., K. K. Divine, D. M. Klinge, F. D. Gilliland, T. Kang, A. G. Schwartz, T. J. Bocklage and S. A. Belinsky (2003). "Promoter hypermethylation of the O6-methylguanine-DNA methyltransferase gene: more common in lung adenocarcinomas from never-smokers than smokers and associated with tumor progression." Cancer Res 63(16): 4842-8.

Raimondi, S., V. Paracchini, H. Autrup, J. M. Barros-Dios, S. Benhamou, P. Boffetta, M. L. Cote, I. A. Dialyna, V. Dolzan, R. Filiberti, S. Garte, A. Hirvonen, K. Husgafvel-Pursiainen, E. N. Imyanitov, I. Kalina, D. Kang, C. Kiyohara, T. Kohno, P. Kremers, Q. Lan, S. London, A. C. Povey, A. Rannug, E. Reszka, A. Risch, M. Romkes, J. Schneider, A. Seow, P. G. Shields, R. C. Sobti, M. Sorensen, M. Spinola, M. R. Spitz, R. C. Strange, I. Stucker, H. Sugimura, J. To-Figueras, S. Tokudome, P. Yang, J. M. Yuan, M. Warholm and E. Taioli (2006). "Meta- and pooled analysis of GSTT1 and lung cancer: a HuGE-GSEC review." Am J Epidemiol 164(11): 1027-42.

Ranson, M., M. R. Middleton, J. Bridgewater, S. M. Lee, M. Dawson, D. Jowle, G. Halbert, S. Waller, H. McGrath, L. Gumbrell, R. S. McElhinney, D. Donnelly, T. B. McMurry and G. P. Margison (2006). "Lomeguatrib, a potent inhibitor of O6-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors." Clin Cancer Res 12(5): 1577-84.

Raunio, H., K. Husgafvel-Pursiainen, S. Anttila, E. Hietanen, A. Hirvonen and O. Pelkonen (1995). "Diagnosis of polymorphisms in carcinogen-activating and inactivating enzymes and cancer susceptibility--a review." Gene 159(1): 113-21.

Raunio, H., A. Rautio, H. Gullsten and O. Pelkonen (2001). "Polymorphisms of CYP2A6 and its practical consequences." Br J Clin Pharmacol 52(4): 357-63.

Ravanat, J. L., T. Douki, P. Duez, E. Gremaud, K. Herbert, T. Hofer, L. Lasserre, C. Saint-Pierre, A. Favier and J. Cadet (2002). "Cellular background level of 8-oxo-7,8-dihydro-2'-deoxyguanosine: an isotope based method to evaluate artefactual oxidation of DNA during its extraction and subsequent work-up." Carcinogenesis 23(11): 1911-8.

Reh, B. D., D. G. DeBord, M. A. Butler, T. M. Reid, C. Mueller and J. M. Fajen (2000). "O(6)-methylguanine DNA adducts associated with occupational nitrosamine exposure." Carcinogenesis 21(1): 29-33.

Rinne, M., D. Caldwell and M. R. Kelley (2004). "Transient adenoviral N-methylpurine DNA glycosylase overexpression imparts chemotherapeutic sensitivity to human breast cancer cells." Mol Cancer Ther 3(8): 955-67.

Rinne, M. L., Y. He, B. F. Pachkowski, J. Nakamura and M. R. Kelley (2005). "N-methylpurine DNA glycosylase overexpression increases alkylation sensitivity by rapidly removing non-toxic 7-methylguanine adducts." Nucleic Acids Res 33(9): 2859-67.

Risch, H. A., G. R. Howe, M. Jain, J. D. Burch, E. J. Holowaty and A. B. Miller (1993). "Are female smokers at higher risk for lung cancer than male smokers? A case-control analysis by histologic type." Am J Epidemiol 138(5): 281-93.

Rojas, M., K. Alexandrov, I. Cascorbi, J. Brockmoller, A. Likhachev, K. Pozharisski, G. Bouvier, G. Auburtin, L. Mayer, A. Kopp-Schneider, I. Roots and H. Bartsch (1998). "High benzo[a]pyrene diol-epoxide DNA adduct levels in lung and blood cells from individuals with combined CYP1A1 MspI/Msp-GSTM1*0/*0 genotypes." Pharmacogenetics 8(2): 109-18.

Rojas, M., I. Cascorbi, K. Alexandrov, E. Kriek, G. Auburtin, L. Mayer, A. Kopp-Schneider, I. Roots and H. Bartsch (2000). "Modulation of benzo[a]pyrene diolepoxide-DNA adduct levels in human white blood cells by CYP1A1, GSTM1 and GSTT1 polymorphism." Carcinogenesis 21(1): 35-41.

Roodi, N., W. D. Dupont, J. H. Moore and F. F. Parl (2004). "Association of homozygous wild-type glutathione S-transferase M1 genotype with increased breast cancer risk." Cancer Res 64(4): 1233-6.

Ruano-Ravina, A., A. Figueiras, L. Loidi and J. M. Barros-Dios (2003). "GSTM1 and GSTT1 polymorphisms, tobacco and risk of lung cancer: a case-control study from Galicia, Spain." Anticancer Res 23(5b): 4333-7.

Rudiger, H. W., U. Schwartz, E. Serrand, M. Stief, T. Krause, D. Nowak, G. Doerjer and G. Lehnert (1989). "Reduced O6-methylguanine repair in fibroblast cultures from patients with lung cancer." Cancer Res 49(20): 5623-6.

Rusin, M., A. Samojedny, C. C. Harris and M. Chorazy (1999). "Novel genetic polymorphisms in DNA repair genes: O(6)-methylguanine-DNA methyltransferase (MGMT) and N-methylpurine-DNA glycosylase (MPG) in lung cancer patients from Poland." Hum Mutat 14(3): 269-70.

Russo, M. T., G. De Luca, P. Degan, E. Parlanti, E. Dogliotti, D. E. Barnes, T. Lindahl, H. Yang, J. H. Miller and M. Bignami (2004). "Accumulation of the oxidative base lesion 8-hydroxyguanine in DNA of tumor-prone mice defective in both the Myh and Ogg1 DNA glycosylases." Cancer Res 64(13): 4411-4.

Sakumi, K., A. Shiraishi, S. Shimizu, T. Tsuzuki, T. Ishikawa and M. Sekiguchi (1997). "Methylnitrosourea-induced tumorigenesis in MGMT gene knockout mice." Cancer Res **57**(12): 2415-8.

Sakumi, K., Y. Tominaga, M. Furuichi, P. Xu, T. Tsuzuki, M. Sekiguchi and Y. Nakabeppu (2003). "Ogg1 knockout-associated lung tumorigenesis and its suppression by Mth1 gene disruption." Cancer Res **63**(5): 902-5.

Sancar, A., L. A. Lindsey-Boltz, K. Unsal-Kacmaz and S. Linn (2004). "Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints." Annu Rev Biochem **73**: 39-85.

Santillan, A. A., C. A. Camargo, Jr. and G. A. Colditz (2003). "A meta-analysis of asthma and risk of lung cancer (United States)." Cancer Causes Control **14**(4): 327-34.

Saparbaev, M., K. Kleibl and J. Laval (1995). "Escherichia coli, Saccharomyces cerevisiae, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA." Nucleic Acids Res **23**(18): 3750-5.

Schabath, M. B., X. Wu, R. Vassilopoulou-Sellin, A. A. Vaporciyan and M. R. Spitz (2004). "Hormone replacement therapy and lung cancer risk: a case-control analysis." Clin Cancer Res **10**(1 Pt 1): 113-23.

Schreiber, G. and D. C. McCrory (2003). "Performance characteristics of different modalities for diagnosis of suspected lung cancer: summary of published evidence." Chest **123**(1 Suppl): 115S-128S.

Sellers, T. A., J. E. Bailey-Wilson, R. C. Elston, A. F. Wilson, G. Z. Elston, W. L. Ooi and H. Rothschild (1990). "Evidence for mendelian inheritance in the pathogenesis of lung cancer." J Natl Cancer Inst **82**(15): 1272-9.

Shapiro, J. A., E. J. Jacobs and M. J. Thun (2000). "Cigar smoking in men and risk of death from tobacco-related cancers." J Natl Cancer Inst **92**(4): 333-7.

Sherratt, P. J., D. J. Pulford, D. J. Harrison, T. Green and J. D. Hayes (1997). "Evidence that human class Theta glutathione S-transferase T1-1 can catalyse the activation of dichloromethane, a liver and lung carcinogen in the mouse. Comparison of the tissue distribution of GST T1-1 with that of classes Alpha, Mu and Pi GST in human." Biochem J **326** (Pt 3): 837-46.

Shields, P. G., A. C. Povey, V. L. Wilson, A. Weston and C. C. Harris (1990). "Combined high-performance liquid chromatography/32P-postlabeling assay of N7-methyldeoxyguanosine." Cancer Res **50**(20): 6580-4.

Shuker, D. E. and G. P. Margison (1997). "Nitrosated glycine derivatives as a potential source of O6-methylguanine in DNA." Cancer Res **57**(3): 366-9.

Sidorkina, O., M. Saparbaev and J. Laval (1997). "Effects of nitrous acid treatment on the survival and mutagenesis of Escherichia coli cells lacking base excision repair (hypoxanthine-DNA glycosylase-ALK A protein) and/or nucleotide excision repair." Mutagenesis **12**(1): 23-8.

Singer, B. and T. P. Brent (1981). "Human lymphoblasts contain DNA glycosylase activity excising N-3 and N-7 methyl and ethyl purines but not O6-alkylguanines or 1-alkyladenines." Proc Natl Acad Sci U S A **78**(2): 856-60.

Singer, B. and B. Hang (1997). "What structural features determine repair enzyme specificity and mechanism in chemically modified DNA?" Chem Res Toxicol **10**(7): 713-32.

Sirzen, F., E. Kjellen, S. Sorenson and E. Cavallin-Stahl (2003). "A systematic overview of radiation therapy effects in non-small cell lung cancer." Acta Oncol **42**(5-6): 493-515.

Sleutels, F., D. P. Barlow and R. Lyle (2000). "The uniqueness of the imprinting mechanism." Curr Opin Genet Dev **10**(2): 229-33.

Slupphaug, G., I. Lettrem, B. Myrnes and H. E. Krokan (1992). "Expression of O6-methylguanine--DNA methyltransferase and uracil--DNA glycosylase in human placentae from smokers and non-smokers." Carcinogenesis **13**(10): 1769-73.

Smith, C. J., T. A. Perfetti, R. Garg and C. Hansch (2003). "IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values." Food Chem Toxicol **41**(6): 807-17.

Smith, G. B., J. R. Bend, L. L. Bedard, K. R. Reid, D. Petsikas and T. E. Massey (2003). "Biotransformation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in peripheral human lung microsomes." Drug Metab Dispos **31**(9): 1134-41.

Smith, L. E., M. F. Denissenko, W. P. Bennett, H. Li, S. Amin, M. Tang and G. P. Pfeifer (2000). "Targeting of lung cancer mutational hotspots by polycyclic aromatic hydrocarbons." J Natl Cancer Inst **92**(10): 803-11.

Smith, S. A. and B. P. Engelward (2000). "In vivo repair of methylation damage in Aag 3-methyladenine DNA glycosylase null mouse cells." Nucleic Acids Res **28**(17): 3294-300.

Smith-Warner, S. A., D. Spiegelman, S. S. Yaun, D. Albanes, W. L. Beeson, P. A. van den Brandt, D. Feskanich, A. R. Folsom, G. E. Fraser, J. L. Freudenheim, E. Giovannucci, R. A. Goldbohm, S. Graham, L. H. Kushi, A. B. Miller, P. Pietinen, T. E. Rohan, F. E. Speizer, W. C. Willett and D. J. Hunter (2003). "Fruits, vegetables and lung cancer: a pooled analysis of cohort studies." Int J Cancer **107**(6): 1001-11.

Sone, S., S. Takashima, F. Li, Z. Yang, T. Honda, Y. Maruyama, M. Hasegawa, T. Yamanda, K. Kubo, K. Hanamura and K. Asakura (1998). "Mass screening for lung cancer with mobile spiral computed tomography scanner." Lancet **351**(9111): 1242-5.

Sorensen, M., H. Autrup, A. Tjonneland, K. Overvad and O. Raaschou-Nielsen (2004). "Glutathione S-transferase T1 null-genotype is associated with an increased risk of lung cancer." Int J Cancer **110**(2): 219-24.

Sorensen, M., O. Raaschou-Nielsen, C. Brasch-Andersen, A. Tjonneland, K. Overvad and H. Autrup (2007). "Interactions between GSTM1, GSTT1 and GSTP1 polymorphisms and smoking and intake of fruit and vegetables in relation to lung cancer." Lung Cancer **55**(2): 137-44.

Sorensen, M., O. Raaschou-Nielsen, R. D. Hansen, A. Tjonneland, K. Overvad and U. Vogel (2006). "Interactions between the OGG1 Ser326Cys polymorphism and intake of fruit and vegetables in relation to lung cancer." Free Radic Res **40**(8): 885-91.

Spitz, M. R., C. M. Duphorne, M. A. Detry, P. C. Pillow, C. I. Amos, L. Lei, M. de Andrade, X. Gu, W. K. Hong and X. Wu (2000). "Dietary intake of isothiocyanates: evidence of a joint effect with glutathione S-transferase polymorphisms in lung cancer risk." Cancer Epidemiol Biomarkers Prev **9**(10): 1017-20.

Stabile, L. P., A. L. Davis, C. T. Gubish, T. M. Hopkins, J. D. Luketich, N. Christie, S. Finkelstein and J. M. Siegfried (2002). "Human non-small cell lung tumors and cells derived from normal lung express both estrogen receptor alpha and beta and show biological responses to estrogen." Cancer Res 62(7): 2141-50.

Staretz, M. E., P. G. Foiles, L. M. Miglietta and S. S. Hecht (1997). "Evidence for an important role of DNA pyridyloxobutylation in rat lung carcinogenesis by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone: effects of dose and phenethyl isothiocyanate." Cancer Res 57(2): 259-66.

Strauss, P. R., W. A. Beard, T. A. Patterson and S. H. Wilson (1997). "Substrate binding by human apurinic/apyrimidinic endonuclease indicates a Briggs-Haldane mechanism." J Biol Chem 272(2): 1302-7.

Su, T., Z. Bao, Q. Y. Zhang, T. J. Smith, J. Y. Hong and X. Ding (2000). "Human cytochrome P450 CYP2A13: predominant expression in the respiratory tract and its high efficiency metabolic activation of a tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone." Cancer Res 60(18): 5074-9.

Sugarbaker, D. J. and G. M. Strauss (2000). "Extent of surgery and survival in early lung carcinoma: implications for overdiagnosis in stage IA nonsmall cell lung carcinoma." Cancer 89(11 Suppl): 2432-7.

Szabo, C., P. Pacher and R. A. Swanson (2006). "Novel modulators of poly(ADP-ribose) polymerase." Trends Pharmacol Sci 27(12): 626-30.

Szyfter, K., K. Hemminki, W. Szyfter, Z. Szmeja, J. Banaszewski and M. Pabiszczak (1996). "Tobacco smoke-associated N7-alkylguanine in DNA of larynx tissue and leucocytes." Carcinogenesis 17(3): 501-6.

Taioli, E. and E. L. Wynder (1994). "Re: Endocrine factors and adenocarcinoma of the lung in women." J Natl Cancer Inst 86(11): 869-70.

Takeuchi, H., K. Saoo, Y. Matsuda, M. Yokohira, K. Yamakawa, Y. Zeng, M. Miyazaki, M. Fujieda, T. Kamataki and K. Imaida (2006). "Dose dependent inhibitory effects of dietary 8-methoxypsoralen on NNK-induced lung tumorigenesis in female A/J mice." Cancer Lett 234(2): 232-8.

Takeuchi, H., K. Saoo, M. Yokohira, M. Ikeda, H. Maeta, M. Miyazaki, H. Yamazaki, T. Kamataki and K. Imaida (2003). "Pretreatment with 8-methoxypsoralen, a potent human CYP2A6 inhibitor, strongly inhibits lung tumorigenesis induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in female A/J mice." Cancer Res 63(22): 7581-3.

Tan, W., G. F. Chen, D. Y. Xing, C. Y. Song, F. F. Kadlubar and D. X. Lin (2001). "Frequency of CYP2A6 gene deletion and its relation to risk of lung and esophageal cancer in the Chinese population." Int J Cancer 95(2): 96-101.

Tang, D., D. H. Phillips, M. Stampfer, L. A. Mooney, Y. Hsu, S. Cho, W. Y. Tsai, J. Ma, K. J. Cole, M. N. She and F. P. Perera (2001). "Association between carcinogen-DNA adducts in white blood cells and lung cancer risk in the physicians health study." Cancer Res 61(18): 6708-12.

Tannenbaum, S. R. (1980). "A model for estimation of human exposure to endogenous N-nitrosodimethylamine." Oncology 37(4): 232-5.

Taverna, P. and B. Sedgwick (1996). "Generation of an endogenous DNA-methylating agent by nitrosation in Escherichia coli." J Bacteriol 178(17): 5105-11.

The Alpha-Tocopherol, B. C. C. P. S. G. (1994). "The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. The Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group." N Engl J Med 330(15): 1029-35.

Thomas, L., L. A. Doyle and M. J. Edelman (2005). "Lung cancer in women: emerging differences in epidemiology, biology, and therapy." Chest 128(1): 370-81.

Thompson, H. J., J. Heimendinger, A. Haegele, S. M. Sedlacek, C. Gillette, C. O'Neill, P. Wolfe and C. Conry (1999). "Effect of increased vegetable and fruit consumption on markers of oxidative cellular damage." Carcinogenesis 20(12): 2261-6.

Thorpe, R., M. Wadhwa and A. Mire-Sluis (1997). "The use of bioassays for the characterisation and control of biological therapeutic products produced by biotechnology." Dev Biol Stand 91: 79-88.

Thun, M. J., S. J. Henley, D. Burns, A. Jemal, T. G. Shanks and E. E. Calle (2006). "Lung cancer death rates in lifelong nonsmokers." J Natl Cancer Inst 98(10): 691-9.

Tockman, M. (1986). "Survival and mortality from lung cancer in a screened population: the Johns Hopkins Study." Chest(89): 325S-26S.

Tokuhata, G. K. and A. M. Lilienfeld (1963). "Familial aggregation of lung cancer in humans." J Natl Cancer Inst 30: 289-312.

Tong, L., M. R. Spitz, J. J. Fueger and C. A. Amos (1996). "Lung carcinoma in former smokers." Cancer 78(5): 1004-10.

Tranah, G. J., J. Bugni, E. Giovannucci, J. Ma, C. Fuchs, L. Hines, L. Samson and D. J. Hunter (2006). "O6-methylguanine-DNA methyltransferase Leu84Phe and Ile143Val polymorphisms and risk of colorectal cancer in the Nurses' Health Study and Physicians' Health Study (United States)." Cancer Causes Control 17(5): 721-31.

Tricker, A. R. (1997). "N-nitroso compounds and man: sources of exposure, endogenous formation and occurrence in body fluids." Eur J Cancer Prev 6(3): 226-68.

Trivedi, R. N., K. H. Almeida, J. L. Fornsaglio, S. Schamus and R. W. Sobol (2005). "The role of base excision repair in the sensitivity and resistance to temozolomide-mediated cell death." Cancer Res 65(14): 6394-400.

Truong, M. T. and R. F. Munden (2003). "Lung cancer screening." Curr Oncol Rep 5(4): 309-12.

Tsuzuki, T., K. Sakumi, A. Shiraishi, H. Kawate, H. Igarashi, T. Iwakuma, Y. Tominaga, S. Zhang, S. Shimizu, T. Ishikawa and et al. (1996). "Targeted disruption of the DNA repair methyltransferase gene renders mice hypersensitive to alkylating agent." Carcinogenesis 17(6): 1215-20.

Turner-Warwick, M., M. Lebowitz, B. Burrows and A. Johnson (1980). "Cryptogenic fibrosing alveolitis and lung cancer." Thorax 35(7): 496-9.

UK Government, O. f. N. S. (2004). Smoking. <u>Living in Britain General Household Survey</u>. **Chapter 8:** 119-50.

UK Government, O. f. N. S. (2006). "Cancer incidence and mortality in the United Kingdom 2001-03."

Vahakangas, K., G. E. Trivers, S. Plummer, R. B. Hayes, H. Krokan, M. Rowe, R. P. Swartz, H. Yeager, Jr. and C. C. Harris (1991). "O(6)-methylguanine-DNA methyltransferase and uracil DNA glycosylase in human broncho-alveolar lavage cells and peripheral blood mononuclear cells from tobacco smokers and non-smokers." <u>Carcinogenesis</u> 12(8): 1389-94.

Varadarajan, S., D. Shah, P. Dande, S. Settles, F. X. Chen, G. Fronza and B. Gold (2003). "DNA damage and cytotoxicity induced by minor groove binding methyl sulfonate esters." <u>Biochemistry</u> 42(48): 14318-27.

Venmans, B. J., T. J. van Boxem, E. F. Smit, P. E. Postmus and T. G. Sutedja (2000). "Outcome of bronchial carcinoma in situ." <u>Chest</u> 117(6): 1572-6.

Vidal, A. E., I. D. Hickson, S. Boiteux and J. P. Radicella (2001). "Mechanism of stimulation of the DNA glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step." <u>Nucleic Acids Res</u> 29(6): 1285-92.

Vincent, R. G., J. W. Pickren, W. W. Lane, I. Bross, H. Takita, L. Houten, A. C. Gutierrez and T. Rzepka (1977). "The changing histopathology of lung cancer: a review of 1682 cases." <u>Cancer</u> 39(4): 1647-55.

Vogel, U., B. A. Nexo, H. Wallin, K. Overvad, A. Tjonneland and O. Raaschou-Nielsen (2004). "No association between base excision repair gene polymorphisms and risk of lung cancer." <u>Biochem Genet</u> 42(11-12): 453-60.

Vulimiri, S. V., X. Wu, W. Baer-Dubowska, M. de Andrade, M. Detry, M. R. Spitz and J. DiGiovanni (2000). "Analysis of aromatic DNA adducts and 7,8-dihydro-8-oxo- 2'-deoxyguanosine in lymphocyte DNA from a case-control study of lung cancer involving minority populations." <u>Mol Carcinog</u> 27(1): 34-46.

Wald, N. J. and H. C. Watt (1997). "Prospective study of effect of switching from cigarettes to pipes or cigars on mortality from three smoking related diseases." <u>Bmj</u> 314(7098): 1860-3.

Wang, D. and W. Sadee (2006). "Searching for polymorphisms that affect gene expression and mRNA processing: example ABCB1 (MDR1)." <u>Aaps J</u> 8(3): E515-20.

Wang, H., W. Tan, B. Hao, X. Miao, G. Zhou, F. He and D. Lin (2003). "Substantial reduction in risk of lung adenocarcinoma associated with genetic polymorphism in CYP2A13, the most active cytochrome P450 for the metabolic activation of tobacco-specific carcinogen NNK." <u>Cancer Res</u> 63(22): 8057-61.

Wang, P., A. B. Guliaev and B. Hang (2006). "Metal inhibition of human N-methylpurine-DNA glycosylase activity in base excision repair." <u>Toxicol Lett</u> 166(3): 237-47.

Waters, T. R., P. Gallinari, J. Jiricny and P. F. Swann (1999). "Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1." <u>J Biol Chem</u> 274(1): 67-74.

Weiserbs, K. F., J. S. Jacobson, M. D. Begg, L. W. Wang, Q. Wang, M. Agrawal, E. P. Norkus, T. L. Young and R. M. Santella (2003). "A cross-sectional study of polycyclic aromatic hydrocarbon-DNA adducts and polymorphism of glutathione S-transferases among heavy smokers by race/ethnicity." Biomarkers 8(2): 142-55.

Weiss, J. M., E. L. Goode, W. C. Ladiges and C. M. Ulrich (2005). "Polymorphic variation in hOGG1 and risk of cancer: a review of the functional and epidemiologic literature." Mol Carcinog 42(3): 127-41.

Weiss, W. (1999). "Asbestosis: a marker for the increased risk of lung cancer among workers exposed to asbestos." Chest 115(2): 536-49.

Wikman, H., A. Risch, F. Klimek, P. Schmezer, B. Spiegelhalder, H. Dienemann, K. Kayser, V. Schulz, P. Drings and H. Bartsch (2000). "hOGG1 polymorphism and loss of heterozygosity (LOH): significance for lung cancer susceptibility in a caucasian population." Int J Cancer 88(6): 932-7.

Wilson, D. M., 3rd (2005). "Ape1 abasic endonuclease activity is regulated by magnesium and potassium concentrations and is robust on alternative DNA structures." J Mol Biol 345(5): 1003-14.

Wolf, P., Y. C. Hu, K. Doffek, D. Sidransky and S. A. Ahrendt (2001). "O(6)-Methylguanine-DNA methyltransferase promoter hypermethylation shifts the p53 mutational spectrum in non-small cell lung cancer." Cancer Res 61(22): 8113-7.

Woodson, K., J. A. Tangrea, M. J. Barrett, J. Virtamo, P. R. Taylor and D. Albanes (1999). "Serum alpha-tocopherol and subsequent risk of lung cancer among male smokers." J Natl Cancer Inst 91(20): 1738-43.

Wu, R. S., S. Hurst-Calderone and K. W. Kohn (1987). "Measurement of O6-alkylguanine-DNA alkyltransferase activity in human cells and tumor tissues by restriction endonuclease inhibition." Cancer Res 47(23): 6229-35.

Wyatt, M. D., J. M. Allan, A. Y. Lau, T. E. Ellenberger and L. D. Samson (1999). "3-methyladenine DNA glycosylases: structure, function, and biological importance." Bioessays 21(8): 668-76.

Xia, L., L. Zheng, H. W. Lee, S. E. Bates, L. Federico, B. Shen and T. R. O'Connor (2005). "Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease." J Mol Biol 346(5): 1259-74.

Xiao, W. and L. Samson (1993). "In vivo evidence for endogenous DNA alkylation damage as a source of spontaneous mutation in eukaryotic cells." Proc Natl Acad Sci U S A 90(6): 2117-21.

Xie, Y., H. Yang, C. Cunanan, K. Okamoto, D. Shibata, J. Pan, D. E. Barnes, T. Lindahl, M. McIlhatton, R. Fishel and J. H. Miller (2004). "Deficiencies in mouse Myh and Ogg1 result in tumor predisposition and G to T mutations in codon 12 of the K-ras oncogene in lung tumors." Cancer Res 64(9): 3096-102.

Xu, C., S. Goodz, E. M. Sellers and R. F. Tyndale (2002). "CYP2A6 genetic variation and potential consequences." Adv Drug Deliv Rev 54(10): 1245-56.

Xu, S., Y. Wang, B. Roe and W. R. Pearson (1998). "Characterization of the human class Mu glutathione S-transferase gene cluster and the GSTM1 deletion." J Biol Chem 273(6): 3517-27.

Yang, M., B. F. Coles, N. E. Caporaso, Y. Choi, N. P. Lang and F. F. Kadlubar (2004). "Lack of association between Caucasian lung cancer risk and O6-methylguanine-DNA methyltransferase-codon 178 genetic polymorphism." Lung Cancer 44(3): 281-6.

Yang, W., J. Y. Lee and M. Nowotny (2006). "Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity." Mol Cell 22(1): 5-13.

Ye, Z., H. Song, J. P. Higgins, P. Pharoah and J. Danesh (2006). "Five glutathione s-transferase gene variants in 23,452 cases of lung cancer and 30,397 controls: meta-analysis of 130 studies." PLoS Med 3(4): e91.

Yu, S. L., S. K. Lee, R. E. Johnson, L. Prakash and S. Prakash (2003). "The stalling of transcription at abasic sites is highly mutagenic." Mol Cell Biol 23(1): 382-8.

Zang, E. A. and E. L. Wynder (1996). "Differences in lung cancer risk between men and women: examination of the evidence." J Natl Cancer Inst 88(3-4): 183-92.

Zhao, C., M. Tyndyk, I. Eide and K. Hemminki (1999). "Endogenous and background DNA adducts by methylating and 2-hydroxyethylating agents." Mutat Res 424(1-2): 117-25.

Zhou, Z. Q., D. Manguino, K. Kewitt, G. W. Intano, C. A. McMahan, D. C. Herbert, M. Hanes, R. Reddick, Y. Ikeno and C. A. Walter (2001). "Spontaneous hepatocellular carcinoma is reduced in transgenic mice overexpressing human O6- methylguanine-DNA methyltransferase." Proc Natl Acad Sci U S A 98(22): 12566-71.

Zienolddiny, S., D. Campa, H. Lind, D. Ryberg, V. Skaug, L. Stangeland, D. H. Phillips, F. Canzian and A. Haugen (2005). "Polymorphisms of DNA repair genes and risk of non-small cell lung cancer." Carcinogenesis.

Zochbauer-Muller, S., K. M. Fong, A. K. Virmani, J. Geradts, A. F. Gazdar and J. D. Minna (2001). "Aberrant promoter methylation of multiple genes in non-small cell lung cancers." Cancer Res 61(1): 249-55.

Appendix One: Smoking, Health and Occupation Questionnaire

Date Completed   ----------------------

Hospital Number   ----------------------

Name   --

Sex   M / F

Date of Birth   -----/-----/-----

Address   --

Post Code   -------------------

Ethnic Group

Bangladeshi		Black Caribbean	
Indian		Black other	
Pakistani		Chinese	
Asian other		White	
Black African		Other	

A.   LIFETIME SMOKING

1. Have you ever smoked on a regular basis?   YES / NO

   (At least 1 cigarette per day for as long as a year)

   If the answer is NO, please go to section C

2. At what age did you start smoking?   _____

3. How many of each of the following do you smoke/have you smoked on average each day throughout your life?

             Packet cigarettes   _____

             Hand rolled cigarettes   _____

             Cigars LARGE / SMALL   _____

             Pipe tobacco   _____

B. CURRENT SMOKING

1. Do you smoke now?   YES / NO

2. If no   when did you stop smoking?

3. If yes when did you last smoke?   _____

    (Smokers – days/hours, ex-smokers – weeks/months)

4. In the last 3 months how many of the following on average have you smoked per day?

    Packet cigarettes   _____

    Hand rolled cigarettes   _____

    Cigars LARGE / SMALL   _____

    Pipe tobacco   _____

5. Have you used nicotine replacement treatment in the last 3 months?   YES/NO

    if yes, what type?_____

    when was the last time you used nicotine replacement?_____

C. PASSIVE SMOKING

1. Are you exposed to passive smoke YES/NO

2. If yes are you exposed   at home

                                  At work

                                  Elsewhere (please specify)

3. In the last 3 months how many hours per day on average have you spent in the presence of people smoking?

    At Home   _____

    At Work   _____

D. ALCOHOL

1. In the last 3 months, on average how many units of alcohol do you drink each week?

    _____ units

E.   MEDICATION

1.   Are you currently taking any medication prescribed by a doctor   YES/NO
2.   If yes what is the medication

3.   Are you currently taking any medicines that you can buy without prescription
      YES/NO
4.   If yes what is the medication?

5.   Are you currently taking any food supplements, vitamins or herbal medications?
      YES/NO
6.   If yes what do you take?

E.   AREA OF RESIDENCE

1.   How long have you lived at the above present address?   _____
      (If less than 1 year go to 2. if more go to 3.)

2.   Where did you live before this?   _____
                                              _____
                                              _____

3.   Do you have a coal fire or paraffin heater in your home or at work that you have used on
      a regular basis over the past 12 months?   YES / NO

F.   OCCUPATION

Now we would like to ask you some questions regarding your work history

1. What are you doing now?

	(Please Tick)
Working full-time	☐
Working part-time	☐
Unemployed but seeking work	☐
Not working because of ill-health/disability	☐
Retired	☐

Other (please specify) ………………………………………………………

2. At what age did you start working? ………………………………………………

If you are currently working please go to question 3

If you are not currently working please go to question 10

If you are currently employed

3. What is the title of your job? (e.g.) …………………………………………………

………………………………………………………………………………………………

4. What type of work do you do? (e.g.) …………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

5. What type of company do you work for? …………………………………………………

………………………………………………………………………………………………

………………………………………………………………………………………………

6. When did you start doing this job? ....

7. Is this a full-time or part-time job? ....

8. In this job are you exposed to any of the following?

	Yes	No
Dust		
Fumes		
Oils		
Plastics/synthetic rubbers		
Asbestos		
Pesticides		
Wood preservatives		
Sealants		
Acids		
Adhesives/glues		
Cleaning agents		
Disinfecting agents		
Ionising radiation or radioactive materials		

9. In this job do you do any of the following

	Yes	No
Welding or gas cutting		
Painting		
Install or remove insulation or fibre panels		
Driving		

Now please go to Q11

If you are no longer working

10. In what year did you stop working? ....

295

11. Please write below the type of work you have done in every job you have held for more than one month since leaving school or college, starting with the last job you held (if no longer working) or the job before the current one?

Job Title	Type of work (what tasks did you do)	Type of company (what did they do?)	Start Month/yr	Finish Month/yr
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				

12 Have you ever been exposed to the following materials in any of your jobs listed in the previous Table?

	Yes	No	If yes please list the job number(s)
Dust			
Fumes			
Oils			
Plastics/synthetic rubbers			
Asbestos			
Pesticides			
Wood preservatives			
Sealants			
Acids			
Adhesives/glues			
Cleaning agents			
Disinfecting agents			
Ionising radiation or radioactive materials			

13. In any of your jobs listed in the Table (Q11) did you do any of the following

	Yes	No	If yes please list the job number(s)
Welding or gas cutting			
Painting			
Install or remove insulation or fibre panels			
Driving			

If we wanted to ask you any further questions about the jobs you have held, can we contact you? Y/N

       If yes what is your phone number?

G. FAMILY HISTORY

1.   Has anyone in your family (direct blood relatives only) been diagnosed with lung cancer?   YES / NO

      If yes give details: _____

2.   Has anyone in your family (direct blood relatives only) been diagnosed with chronic bronchitis or emphysema?   YES / NO

      If yes give details: _____

3.   Has anyone in your family (direct blood relatives only) been diagnosed with any other lung / respiratory condition?   YES / NO

      If yes give details: _____

H. PAST MEDICAL HISTORY

1.   Have you ever been diagnosed with lung cancer?   YES / NO

      If yes when was this diagnosis?   _____

      What treatment have you received?   _____

2.   Have you ever been diagnosed with chronic bronchitis or emphysema?

                                        YES / NO

      If yes when was this diagnosis made?   _____

3.   Have you ever been diagnosed with asthma?   YES / NO

      If yes when was this diagnosis made?   _____

4.   Have you ever been diagnosed with any other lung / respiratory condition?

                                        YES / NO

      If yes give details: _____

Appendix Two: Patient Information Leaflet (Case-Control Study)

Study Title: **Determinants of Lung Cancer Susceptibility**

Introduction

You are being invited to take part in a research project which is being undertaken by the Bronchoscopy Unit at Wythenshawe Hospital (supervisor Dr P Barber, Consultant Chest Physician) in collaboration with Manchester University (Dr AC Povey, Centre for Occupational and Environmental Health).
Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

We want to investigate several different genetic / chemical markers that could identify people at increased risk of developing lung cancer. This will help to develop future screening strategies to try to diagnose lung cancer at an early stage and so improve treatment success.

Who can take part in the study?

We are asking all people over the age of 18 who are due to undergo a bronchoscopy, for whatever reason, to participate in this study. It is very important that as many people as possible take part.

Do I have to take part?

No. It is entirely up to you to decide whether or not to take part. If you decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.
Your decision will not affect the care you receive at all.

What will happen to me if I take part?

Your involvement in the study is limited to the day of your bronchoscopy only. Your participation will involve three aspects:

1) We will ask you to complete a short questionnaire regarding your smoking habits, and working history. This will take approximately fifteen minutes prior to the procedure.
2) We would like to take a single blood sample to assess levels of smoking related chemicals and extract genetic material from white blood cells for further testing. This will be no different to a normal blood test.
3) We would like to collect a small number of cells from your lung during the bronchoscopy procedure. To do this we need to perform an extra test called a bronchial lavage. This involves:

    - flushing an eggcup of salty water into the right upper part of your lung
    - then immediately sucking this water back into a pot for collection.

This will only be done if you are tolerating the procedure well. It will only add one to two minutes to the procedure as a whole.

What are the risks/ disadvantages of taking part in the study?

The disadvantages are that asking questions may be regarded as an intrusion and would take about 15 minutes of your time. You may also be concerned about the biological samples being taken. The bronchial lavage is not associated with any significant additional side effects. On a rare occasion wheezing can occur; this generally settles within an hour of the procedure.

The usual risks of bronchoscopy would be unchanged from normal.

What are the benefits of taking part in the study?

There are no direct benefits to you from the study. However you would be helping the research team to understand better how lung disease develops and whether there are certain people who may be more susceptible to developing the disease. This may be very important in helping us to develop strategies for early lung cancer detection.

Confidentiality

Any information obtained from the questionnaire or samples taken will be treated with full medical confidentiality by the researchers involved. Your information will not be released to any other third party without your written consent.

Consent

If you give your consent to take part in the study this will not affect your medical care or legal rights. Your participation is voluntary and you are free to withdraw at any time. By giving your consent you are agreeing that you understand what the aims of the study are and what the study requires.

Who is funding the study?

The study is being funded by the NHS and the Colt Foundation, a charity that is particularly interested in identifying the cause of illnesses arising from conditions at the place of work. The people carrying out the study receive no personal benefit and get a standard University/NHS salary for the work that is carried out.

PATIENT CONSENT FORM

Study Title: **Determinants of Lung Cancer Susceptibility**

I ……………………………………… (Patient) have read the explanation of the study in this leaflet and am satisfied with the answers to questions I have asked. I agree that I will take part in the study. I understand that I may withdraw consent at any time without prejudice to my treatment. This study does not alter my statutory rights.

Signature…………………………………

Date………………………………………

I ……………………………………… (Investigator) confirm that the patient named above has read and apparently understood the information provided about the study. I have fully answered any questions/queries they have had and explained all information from this study will be handled with full medical confidentiality.

Signature…………………………………

Date………………………………………

Thank you very much for your participation in this study.

<div align="right">
Dr Philip Crosbie
Respiratory Research Registrar,
North West Lung Centre,
Tel. 0161 291 2721
</div>