CASE-ONLY STUDY OF INTERACTIONS BETWEEN SPECIFIC GENETIC POLYMORPHISMS AND CIGARETTE SMOKING IN THE AETIOLOGY OF PARKINSON’S DISEASE

BY

Yifu Deng

Bachelor of Medicine (Fudan University Medical School, PRC)
Master of Medicine (Beijing University Medical School, PRC)

A thesis submitted for the Degree of Doctor of Philosophy in the Centre for Health Research, Queensland University of Technology

August 2005
ABSTRACT

The aetiology of Parkinson’s disease (PD) is still unclear. Research findings suggest that both environmental and genetic factors may contribute to its development. The interactions between genes and the environment might exist and play a key role. Cigarette smoking was found to be one of the few factors exhibiting a protective effect. If chemical compounds found in cigarette smoke influence PD risk, the difference in the ability of certain individuals in metabolising these substances might alter their susceptibility to the risk of developing PD. Many metabolic enzyme genes exhibit polymorphic traits with alteration of gene function. These might be associated with an altered susceptibility of individuals to PD. Few studies have examined the hypothesis that metabolic enzyme gene polymorphisms might modulate the effect of smoking on PD risk. However, it is crucial to consider these potential interactions when we try to elucidate the aetiology of PD. Even if each factor only contributes a slight variation and influences a small portion of the whole population, non-linear and unpredictable interactions may account for a high proportion of the aetiological fraction. Previous studies have not been strictly designed to examine the interactions between smoking and metabolic enzyme genetic polymorphisms. These studies have not been able to elucidate the extent of the interaction. Therefore, this PhD project attempted to examine whether genetic factors, operating in the phase one and phase two metabolic pathways, interact with smoking to influence the development of PD. This is the first genetic epidemiological study of PD specifically addressing this issue. The research aids in further understanding the aetiology of PD and may be useful for identifying people at higher risk.
A case-only design was chosen for this project for two reasons: first, PD is a relatively rare disease and the case-only design is much more efficient at detecting gene-environment interactions; second, the PD cases for the project were recruited over the past few years and represent a prevalence series, for which an appropriate comparison group for the cases is difficult to identify and recruit. In a case-only study, only cases are used to investigate the multiplicative effects of the exposures and susceptible genotypes of interest, while non-case subjects (traditionally controls) are solely used to test the independence between the exposure and the susceptible genotype. Therefore, this approach avoids the challenges of control selection, a major limitation inherent in the case-control approach.

This thesis comprised of three independent studies: the first study investigated the interactions between genetic polymorphisms of GSTM1, P1, T1 and Z1 and smoking in PD; the second study examined the interactions between genetic polymorphisms of CYP2E1 and smoking in PD; and the third study examined the interactions between genetic polymorphisms of CYP2D6 and smoking in PD. The first two studies recruited 400 white Caucasian PD cases from both hospital wards and private neurology clinics (230 men and 170 women). The third study further included 142 white Caucasian PD cases newly recruited from the same sources (542 in total, 321 men, and 221 women). The mean age of cases was 67 years with the average onset age at 60 years.

GSTM1, GSTP1, GSTT1, GSTZ1 AND CYP2E1 genotyping processes were performed using protocols previously published with minor modification, whereas
CYP2D6 genotyping methods were mainly developed by me with assistance from associate supervisor Dr. George Mellick. Reliability and validity of the PCR and RFLP methods were assessed through re-conducting the genotype assays using at least a 10% sample of our DNA samples. The results for all re-assessments were 100% concordant.

Crude bivariate analyses were adjusted for potential confounding effects of the variables, including age at onset, gender, family history of PD and pesticide exposures. Among our unaffected, aged subjects (mean age: 63.9 years, sd: 11.4 years), the genotype frequencies at each locus were similar to those reported in other Caucasian populations. The first study showed that the proportion of carriers of the GSTP1-114Val allele (mutant) increased with increasing smoking dose from 0 to >30 pack-years. Homozygotes of the 114Ala allele (wild-type) decreased with increasing smoking dose (trend test: p=0.02). This trend existed both in male and female cases. This dose-effect relationship was most significant in the group of cases with late-onset PD (i.e., age at onset > 55 years) with the ORi\textsubscript{case-only} values of 1.88 (95%CI: 0.65-5.48) and 2.63 (95%CI: 1.07-6.49) for >0-10 and >10 pack-years, respectively. No similar trend was found among our unaffected, aged subjects (p=0.42). Haplotype analyses revealed significant differences for GSTP1 haplotypes between smoking and non-smoking PD cases (OR\textsubscript{case-only} for *C haplotype=2.00 (95%CI: 1.11-3.60), p=0.03). In this case, smoking-exposed PD cases were more likely to possess the *C haplotype defined by A to G base-pair transition at nucleotide +313 and C to T base-pair transition at nucleotide +341 (at amino acid level, valine at both positions 105 and 114). The second study found no difference in CYP2E1 genotype frequencies between PD cases who ever smoked compared to those who never smoked (odds ratio
for interaction (ORi) = 1.00 (95% CI: 0.39-2.51, p=0.99)). No CYP2E1 gene-smoking interactions were detected in relation to age at onset of PD. The third study found that among cases without regular pesticide exposures, CYP2D6 PMs who smoked more than 5 pack-years had a later mean age at disease onset (68.6 years) than those with extensive metaboliser phenotypes (EMs) (61.1 years, p=0.02) and non-smokers (60.5 years, p=0.01). Analysis of aged subjects without PD confirmed that neither smoking status nor CYP2D6 PM status was associated with age itself.

Our data suggest: 1. smoking exposure is independent of GSTM1, P1, T1, Z1 and CYP2E1 genotypes; 2. smoking may be, to some extent, associated with CYP2D6 genotypes; 3. there are no multiplicative interactive effects linking smoking and GSTM1, T1, Z1 or CYP2E1 genotypes with the risk for PD; 4. there is a multiplicative interactive effect between smoking and GSTP1 haplotype – particularly for genotypes carrying the 114Val allele; and 5. there is a multiplicative interactive effect between smoking and CYP2D6 PMs – particularly for people who ever smoked cigarettes more than 5 pack-years.

In general, this thesis provides a model for exploring the gene-smoking interactions in PD. Further studies need to consider the recruitment of a large number of population-based and randomly-selected samples and to pay more attention to measurement of environmental exposures. Further studies also need to examine simultaneously the impact of smoking, pesticide exposures and other potential risk factors on PD. These studies will build evidence for interactions contributing to this common neurological movement disorder.
KEYWORDS: Parkinson’s disease, smoking, GSTM1, GSTP1, GSTT1, GSTZ1, CYP2D6, CYP2E1, gene-environment interaction.
PUBLICATIONS RELEVANT TO THE THESIS

JOURNAL ARTICLES

Deng YF, Newman B, Dunne MP, Silburn PA and Mellick GD: Case-only study of interactions between genetic polymorphisms of GSTM1, P1, T1 and Z1 and smoking in Parkinson’s disease. Neuroscience Letters. 366(3):326-331, 2004


MANUSCRIPT

Deng YF, Newman B, Dunne MP, Silburn PA, Mellick GD: No interactive effect between the CYP2E1 genetic polymorphisms and smoking in Parkinson’s disease: a case-only study

Deng YF, Gartner C, Newman B, Dunne MP, Silburn PA, Boyle RS, Mellick GD: CYP2D6 poor metabolisers who smoke cigarettes show a later onset of Parkinson’s disease

CONFERENCE PRESENTATIONS

Deng YF, Newman B, Dunne MP, Silburn PA, Mellick GD (2004): Case-only study of interactions between genetic polymorphisms of GSTM1, P1, T1 and Z1 and smoking in Parkinson’s disease.

- Poster presentation: The 15th International Conference on Parkinson’s Disease, Beijing, China


- Oral presentation: The 3rd National Emerging Researchers in Aging (ERA) Conference, Brisbane, Australia


- Oral presentation: The 3rd National Emerging Researchers in Aging (ERA) Conference, Brisbane, Australia


STATEMENT OF AUTHORSHIP

The work contained in this thesis has not been previously submitted for a degree or diploma at any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signed: ______________________________
Date: ________________________________
ACKNOWLEDGEMENTS

I would like to thank many people who had an important role in the completion of this thesis.

First of all, the greatest thanks and appreciation to my supervisors, Associate Professor Michael P. Dunne, Professor Beth Newman, and Dr. George D. Mellick, for their capable and experienced professional guidance. At all times throughout my candidature, they have maintained diligence, interest and enthusiasm for my research. As a student from non-English speaking background, I have been very lucky to be generously instructed by my supervisor team, not only academically, but also linguistically and culturally. To all my supervisors, I would like to thank for their significant amount of time spent on their professional guidance on my completing the thesis. I am grateful to them for their capacity as experienced researchers in looking over my project.

I am grateful to my wife, Jianmin Ding and my daughter, Mengmei Deng, for their patience, encouragement and emotional support to me throughout.

Special thanks to Dr. Diana Battistatta for her statistical advice and helpful comments on my earlier manuscripts.

I am also grateful to Ms. Coral Gartner who spent her time in assistance with my data collection.

A thank you also to the many people who assisted with proof-reading my language expression.

I would also like to acknowledge all my colleagues in the Centre for Health Research for their advice and assistance with research and personal friendship.
## TABLE OF CONTENTS

### CHAPTER 1: INTRODUCTION AND BACKGROUND

1.1 PREFACE ........................................................................................................... 1
1.2 PATHOGENESIS OF PD ..................................................................................... 2
1.3 RISK FACTORS CONTRIBUTING TO PD .......................................................... 2
1.4 SMOKING, SPECIFIC GENETIC POLYMORPHISMS AND PD ......................... 4
1.5 CONSIDERATION OF STUDY DESIGN ........................................................... 6
1.6 PROJECT AIM, OBJECTIVES AND HYPOTHESES .......................................... 7
1.7 CONTENTS AND STRUCTURE OF THE THESIS ............................................ 7
1.8 SIGNIFICANCE OF THE THESIS .................................................................... 9

### CHAPTER 2: LITERATURE REVIEW: CIGARETTE SMOKING, SPECIFIC
GENETIC POLYMORPHISMS OF METABOLIC ENZYMES AND THE RISK
OF PARKINSON’S DISEASE .................................................................................... 13

2.1 INTRODUCTION .............................................................................................. 13
2.2 CIGARETTE SMOKING AND PD RISK ............................................................ 14
2.3 GENETIC INFLUENCE ON CIGARETTE SMOKING ......................................... 23
2.4 HUMAN GENETIC POLYMORPHISMS OF METABOLIC ENZYMES AND PD RISK .......................................................... 26
2.5 GENETIC POLYMORPHISMS OF METABOLIC ENZYMES AND
CIGARETTE SMOKING IN COMBINATION AS A RISK MODIFIER OF PD ............. 45
2.6 CONCLUSIONS ............................................................................................... 46

### CHAPTER 3: METHODOLOGY ............................................................................ 54

3.1 INTRODUCTION .............................................................................................. 54
3.2 RESEARCH DESIGN ....................................................................................... 54
3.3 PD CASES ...................................................................................................... 59
3.4 UNAFFECTED SUBJECTS ................................................................................ 62
3.5 DATA COLLECTION .......................................................................................... 66
3.6 CANDIDATE GENES ...................................................................................... 67
3.7 DNA EXTRACTION ......................................................................................... 68
3.8 OPTIMISATION OF GENOTYPING ASSAYS ..................................................... 68
CHAPTER 4: CASE-ONLY STUDY OF INTERACTIONS BETWEEN GENETIC POLYMORPHISMS OF GSTM1, P1, T1, Z1 AND SMOKING IN PARKINSON’S DISEASE

ABSTRACT
4.1 INTRODUCTION
4.2 MATERIALS AND METHODS
4.3 RESULTS
4.4 DISCUSSION
ACKNOWLEDGEMENTS
ATTACHMENT

CHAPTER 5: NO EVIDENT INTERACTIONS BETWEEN THE CYP2E1 GENETIC POLYMORPHISM AND CIGARETTE SMOKING IN PARKINSON’S DISEASE: A CASE-ONLY STUDY

ABSTRACT
5.1 INTRODUCTION
5.2 MATERIALS AND METHODS
5.3 RESULTS
5.4 DISCUSSION
ACKNOWLEDGEMENTS
ATTACHMENT

CHAPTER 6: CYP2D6 POOR METABOLIZERS WHO SMOKE CIGARETTE SHOW A LATER AGE AT ONSET OF PARKINSON’S DISEASE

ABSTRACT
6.1 INTRODUCTION
6.2 MATERIALS AND METHODS
LIST OF TABLES

TABLE 2.1: REVIEW OF STUDIES INVESTIGATING THE ASSOCIATION BETWEEN PD AND VARIOUS GST ENZYME GENES 49-50

TABLE 2.2: REVIEW OF STUDIES INVESTIGATING THE ASSOCIATION BETWEEN PD AND CYP2D6 ENZYME GENES 51-52

TABLE 2.3: REVIEW OF STUDIES INVESTIGATING THE ASSOCIATION BETWEEN PD AND CYP2E1 ENZYME GENES 53

TABLE 3.1: GENE-ENVIRONMENTAL INTERACTION DEFINITION 55

TABLE 3.2: GENE-ENVIRONMENT INTERACTION IN RELATION TO AGE AT ONSET OF PD 58

TABLE 3.3: DEMOGRAPHIC CHARACTERISTICS OF UNAFFECTED AGED SUBJECTS 62

TABLE 3.4: REQUIRED SAMPLE SIZES 65

TABLE 3.5: SUMMARY OF VARIABLES FOR INTERACTION STUDY 71

TABLE 3.6: SUMMARY OF VARIABLES FOR INDEPENDENCE OF SMOKING AND GENOTYPE 72

TABLE 4.1: CHARACTERISTICS OF PD CASES 92

TABLE 4.2: GENOTYPE FREQUENCIES AMONG NEVER- AND EVER-SMOKERS IN PD CASES 93

TABLE 4.3: PACK-YEARS AND GSTP1-114 GENOTYPES AMONG PD CASES 94

TABLE A-4.1: PRIMER SEQUENCES AND ENDONUCLEASES 95

TABLE A-4.2: INDEPENDENCE OF SMOKING STATUS AND GENOTYPES IN UNAFFECTED INDIVIDUALS 96

TABLE A-4.3: AGE AT ONSET, PACK-YEARS AND GSTP1-114 GENOTYPES AMONG PD CASES 97

TABLE A-4.4: HAPLOTYPE FREQUENCIES AMONG NEVER- AND EVER-SMOKERS IN PD CASES 98

TABLE A-4.5: AGE AT ONSET OF PD AMONG EVER- AND NEVER-SMOKERS 99
**TABLE 5.1:** DEMOGRAPHIC CHARACTERISTICS OF PD CASES  
**TABLE 5.2:** CYP2E1 GENOTYPE AND ALLELE FREQUENCIES AMONG NEVER- AND EVER-SMOKERS IN PD CASES  
**TABLE 5.3:** AGE AT ONSET OF PD AMONG EVER- AND NEVER-SMOKERS  
**TABLE A-5.1:** CYP2E1 GENOTYPES AND SMOKING STATUS IN UNAFFECTED INDIVIDUALS  
**TABLE 6.1:** CHARACTERISTICS OF THE CASES  
**TABLE 6.2:** CYP2D6 GENOTYPES, SMOKING AND AGE AT ONSET OF PD  
**TABLE A-6.1:** CYP2D6 GENOTYPES AND SMOKING STATUS  
**TABLE A-6.2:** CYP2D6 GENOTYPES AND SMOKING HABIT  
**TABLE A-6.3:** PESTICIDE EXPOSURES AND AGE AT ONSET
LIST OF FIGURES

FIGURE A-4.1: MULTIPLEX PCR ASSAY FOR THE PRESENCE OF GSTM1 AND GSTT1 GENES 100
FIGURE A-4.2: PCR-RFLP ANALYSIS OF GSTP1 GENOTYPES 101
FIGURE A-4.3: PCR-RFLP ANALYSIS OF GSTP1 GENOTYPES 102
FIGURE A-6.1: PCR-RFLP ANALYSIS OF CYP2D6 GENOTYPES 131
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%CI</td>
<td>95% confidential intervals</td>
</tr>
<tr>
<td>CYP</td>
<td>Cytochrome p450</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione s-transferase</td>
</tr>
<tr>
<td>NHMRC</td>
<td>National Health and Medical Research Council</td>
</tr>
<tr>
<td>PAH</td>
<td>The Princess Alexandra Hospital</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>QUT</td>
<td>Queensland University of Technology</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphism</td>
</tr>
<tr>
<td>UQ</td>
<td>University of Queensland</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION AND BACKGROUND

1.1 PREFACE

Parkinson’s Disease (PD) is a neurodegenerative disease leading to a movement disorder characterised by symptoms including bradykinesia (slowness of movement), muscular rigidity, resting tremor and postural instability. It is often accompanied by cognitive and affective disturbances, sensory symptoms, and sleep disturbance. Preliminary observations indicate that PD is not evenly distributed across the world. The prevalence of PD in Caucasian populations has been reported to be between 150 and 300 per 100,000 [1]. Although PD is relatively rare in the whole population, the prevalence of PD rises sharply with age and is common in the elderly. A study by Bennett and colleagues indicated that parkinsonian signs (of which PD is the main cause) were observed in 15% of people aged 65 to 74 years, 30% of people 75 to 84 years and 50% of people 85 years and older [2]. In Australia, a more recent study by Chan and colleagues demonstrated that the crude prevalence of PD was between 3.6 and 4.9% for people aged 55 and over, and the age-adjusted prevalence of PD has increased at least 42.5%, compared to the prevalence published in 1966 [3]. Worldwide populations, particularly in developed countries like Australia, are aging. Therefore, PD has become a significant public health concern. Furthermore, there is no specific cure for PD at present and the aetiology of PD is still unclear.
1.2 PATHOGENESIS OF PD

The major neuropathological alteration in PD is significant loss of the dopamine-containing neurons in the compact zone of the substantia nigra, caudate nucleus and putamen (striatum) in the brain. Moreover, significant reduction of dopamine concentration can be detected in these areas. A further morphologic hallmark of PD is the presence of Lewy bodies in the remaining neurons. Various mechanisms have been implicated in the pathogenesis and progression of PD. Among them, excessive formation and/or lack of detoxification of destructive oxygen radicals and hydrogen peroxide in critical areas of the brain have been suggested to play an important role [4, 5]. Similarly, dysfunction in mitochondrial respiratory enzymes could be a key event [6]. The primary events that induce mitochondrial failure and oxidative damage are unknown. However, these abnormalities may trigger apoptotic cell death, an explanation for the neuronal loss in PD. More recent studies have indicated that the common final process of neuro-degeneration could be failure of the ubiquitin-proteasome system. It has been noted that the substantia nigra pars compacta of PD patients accumulates high amounts of oxidised and nitrated proteins that are resistant to proteasomal degradation [7-10]. Lewy bodies also contain various proteins that should have been degraded in normal situations [9, 10]. However, the exact mechanism by which abnormal protein accumulation leads to neuronal loss is unclear.

1.3 RISK FACTORS CONTRIBUTING TO PD

PD risk increases with aging. Pathologically, aging has been proven to be associated with a decline of dopamine-containing neurons in the compact zone of the substantia
nigra. However, accumulated evidence suggests that PD is unlikely to be simply an acceleration of aging.

For several decades, researchers have noticed that PD might be associated with environmental factors [11-17]. However, large-scale investigations of potential environmental risk factors for PD were not triggered until the chance observation that parkinsonian signs and symptoms can be induced by a drug known as 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) in 1983 [18]. The toxic substance involved was shown to be a monoamine-oxidase-derived metabolite of MPTP named 1-methyl-4-phenyl pyridium ion (MPP+) [19, 20]. MPP+ selectively targets dopaminergic cells and inhibits mitochondrial complex 1 in the respiratory chain [21, 22]. MPP+ is of particular interest since it is chemically related to the well-known herbicide Paraquat [23]. Experimentally, when Paraquat was given to frogs in the presence of cytochrome P450, it presented the same toxic effect as MPP+, selectively depigmenting the skin melanised areas [24]. These observations suggested that environmental toxicants might cause PD [25-40]. Further epidemiological studies examining environmental factors demonstrated that exposures to pesticides, metals and organic solvents, head injury, rural living as well as well-water consumption were associated with an increased risk of PD [41-53]. Interestingly, cigarette smoking and tea and coffee consumption appeared to decrease PD risk [54-60].

Besides environmental factors, it was also consistently noticed that PD patients have a positive family history more often than unaffected individuals, indicating that genetic risk factors for the disease might also exist [61, 62]. Twin studies, using the technique of positron emission tomography (PET) that can detect reduction in dopamine uptake
prior to symptomatic disease, also indicated that genetic factors might be associated with susceptibility to PD [63]. Rare mutations in genes such as alpha-synuclein [64, 65], ubiquitin carboxy-terminal hydrolase L1 (UCHL1) [66], parkin [67-69] and tau [70] may result in cases of familial PD with autosomal dominant or recessive mode of inheritance. However, this direct link between genetic mutation and PD only accounts for a very small portion of PD [71, 72]. In order to investigate the contribution of genetic factors to idiopathic PD, many studies, using a case-control design, have examined associations between low-penetrance genes, such as metabolic enzyme genes and PD risk [73-90]. The results are notoriously unreliable but suggest that some polymorphisms in a variety of genes might influence the susceptibility of individuals to PD. However, to date, no specific genetic factors have been reproducibly confirmed for the common form of idiopathic PD.

It has been commonly accepted that, for most idiopathic PD cases, the aetiology is likely to be multi-factorial. Each genetic and environmental factor contributes a fraction to the overall risk. More challengingly, recent evidence has strongly indicated that interactions between genetic and environmental factors might exist and these interactions may play a key role in the aetiology of idiopathic PD [91-97].

1.4 SMOKING, SPECIFIC GENETIC POLYMORPHISMS AND PD

Cigarette smoking is one of the few factors that consistently appear to alter risk for the disease, exhibiting an inverse association [56-60]. Our analyses have also confirmed the protective effect of cigarette smoking on PD with an OR of 0.6 (95% confidence interval: 0.4-0.8) (unpublished). However, the underlying mechanism is still unclear.
Some researchers have suggested that the potential mechanism of smoking’s protective effect might come from: (1) a chemical and/or biochemical process, whereby substances in cigarette smoke exert a direct or indirect protective effect on nigral neurons or glia to reduce the damage to them [60, 98-102]; and/or, (2) an alteration in the metabolism of endogenous neurochemical compounds or exogenous substances due to changes in metabolic enzyme activities or by competition between these substrates and other constituents of cigarette smoke, which may inactivate otherwise neurotoxic substances [98-100, 102].

Theoretically, if chemical compounds found in cigarette smoke influence PD risk, the difference in the ability of certain individuals to metabolise these substances may alter their susceptibility to developing PD. Many metabolic enzyme genes have been proven to exhibit polymorphic traits and some of these genetic variations directly alter gene function. However, few previous studies have examined this hypothesis.

This PhD thesis focuses on examining the potential interactions between cigarette smoking and specific polymorphisms of the metabolic enzyme genes of GSTM1, GSTP1, GSTT1, GSTZ1, CYP2D6, and CYP2E1. This is justified for three reasons: first, sequencing of these metabolic enzyme genes revealed that they exhibit polymorphic traits that might alter gene function; second, all these proven polymorphisms have relatively high allelic frequencies (greater than 5%); and third, they are involved in the metabolism of chemical compounds in cigarette smoke.
1.5 CONSIDERATION OF STUDY DESIGN

In general, the classical case-control approach has been commonly adopted in association studies for rare and late-onset diseases such as PD. Despite the numerous advantages, this approach has its inherent limitations, particularly in recruiting the appropriate comparison group (i.e., control subjects). To date, few previous studies have been strictly designed to examine interactions between cigarette smoking and metabolic enzyme genetic polymorphisms in relation to PD risk. Therefore, these studies were unable to elucidate the extent of the potential interactions.

This PhD project applied the case-only design, which is presently considered the most economical and powerful method for studying gene-environment interactions for susceptibility genes [103]. In this method, only cases are used to investigate the multiplicative effects of the exposures and susceptible genotypes of interest, while non-case subjects (traditionally controls) are used to test the independence between the exposure and the susceptible genotype (see Chapter 3) [103]. Compared to the traditional case-control design, a case-only design requires smaller sample size to detect the same magnitude of interactive effects [103, 104]. Clearly this is very important when we consider that in most situations, the interactions are only subtle. Further, the utilisation of the case-only design also minimises many of the limitations inherent in the case-control approach. In particular, the challenges of control selection are avoided.
1.6 PROJECT AIM, OBJECTIVES AND HYPOTHESES

The main aim of this project is to examine whether specific genetic factors operating in the phase one and phase two metabolic pathways interact with cigarette smoking to influence the development of PD. The three objectives are:

- to examine whether specific polymorphic alleles are correlated with smoking status or exposure dose;
- to identify whether cigarette smoking exposure interacts with specific genetic factors to influence PD; and
- to evaluate the extent of the potential interactions between cigarette smoking and specific polymorphic alleles in the aetiology of PD risk.

We hypothesized that polymorphic alleles at selected loci in the GSTM1, GSTT1, GSTP1, GSTZ1, CYP2D6 and CYP2E1 genes are not linked to smoking status or exposure dose, and that polymorphic alleles at these selected loci may interact with cigarette smoking to influence the risk of developing PD.

1.7 CONTENTS AND STRUCTURE OF THE THESIS

This thesis is presented in publication style. As such, it contains three manuscripts, each designed to stand on its own. Following this Chapter, Chapter 2 critically reviews the literature relating to the association studies of cigarette smoking, targeted specific genetic polymorphisms and PD risk.
Chapter 3 provides an overview of the methods for three studies based on the case-only design. The issues discussed in Chapter 3 are those not dealt with (or only discussed in minimum detail) in the Methods sections of each of the manuscripts. Chapter 3 also addresses the rationale for selecting a case-only design instead of a traditional case-control design for the three studies.

The three manuscripts are presented in Chapters 4 through 6. Each manuscript was written in the conventional publication style for the journal to which it was/is targeted. However, the referencing style of each manuscript has been altered to be consistent with that of the thesis. Each manuscript addresses the objectives of the overall study in terms of the specific results and discussion included. As each manuscript is written independently for the journals, there is some repetitiveness in their introduction, methods and discussion sections.

The first manuscript (Chapter 4) investigates the interactions between genetic polymorphisms of GSTM1, P1, T1 and Z1 and smoking in PD. This manuscript has been published in Neuroscience Letters (2004). The second manuscript (Chapter 5) examines the interactions between genetic polymorphisms of CYP2E1 and smoking in PD, whereas the third manuscript (Chapter 6) examines the interactions between genetic polymorphisms of CYP2D6 and smoking in PD. As stated above, all three manuscripts were based on a case-only approach.

The final Chapter (Chapter 7) summarises the study findings across the three manuscripts and discusses conclusions in relation to the overall aims of the study. In
addition, this Chapter further discusses the studies’ limitations, directions for future research, and the public health implications of the findings.

Tables and figures are provided at the end of each chapter. The references for each of the manuscripts are presented at the end of their corresponding chapters. A complete reference list (including references cited in the manuscripts) is provided at the end of the thesis.

1.8 SIGNIFICANCE OF THE THESIS

This PhD research project directly focused on genes that are involved with the metabolism of constituents of cigarette smoke. As far as we know, this is the first genetic epidemiological study of PD specifically addressing this issue. The findings of this thesis are suggestive for future studies of PD as well as other neurodegenerative diseases. These findings may also be generalised to wider public health and medical conditions. Significance of this PhD project is summarised in the following points:

- **PD is a significant public health and medical issue**

  With populations aging, developed countries like Australia are facing, more than ever, age-related health problems. In response to this trend, in 1995 the Australian National Health and Medical Research Council (NHMRC) and the Australian Science and Technology Council listed aging and aging-related diseases, including aging-related neurodegenerative disorders, as a priority health issue for Australia in the 21st century. PD, the most common serious movement disorder and the second most common aging-related neurodegenerative disorder (following Alzheimer's disease), has already
shown both increased incidence and increased prevalence in the Australian population. It has become a major public health and medical concern in Australia. Therefore, this research and outcomes of this thesis address priority health issues in Australia.

- **Better understanding of the risk factors of PD**

Understanding the risk factors contributing to PD could be crucial in the battle against this disease. However, the aetiology of PD is still far from clear. In 1999, the USA National Institutes of Health, National Institute of Environmental Health Sciences and National Institute of Neurological Disorders and Stroke started to provide specific funding for research aimed at elucidating the role of the environment in PD. Similarly, in 1998, the Australian NHMRC Wills Report stated that identifying genes and environmental causes of PD is one of seven priority areas that require more research. This PhD thesis focused on investigating the interactions between an environmental factor (cigarette smoking) and genetic factors (specific genetic polymorphisms of the major metabolic enzyme genes). Hence, this research project is responsive to national and international priorities.

- **Better understanding of gene-environment interactions in PD**

Although gene-environment interactions have been suggested to play a key role in the aetiology of idiopathic PD, to date there have been very limited attempts to examine the influence of genetic factors on the effect of environmental exposures, or alternatively, the influence of environmental exposures on the effect of genetic factors. The main underlying reason is that such examination is difficult since it is generally unclear what exposures may influence PD risk and, moreover, how these exposures should be assessed. In principle, the ability to examine gene-environment interactions
requires an accurate assessment of the environmental exposures. Since cigarette smoking has been proven to be one of the very few assessable exposures influencing PD risk, studying gene-smoking interactions in the aetiology of PD could provide an excellent model to investigate other gene-environment interactions in the alteration of PD risk.

- **Better understanding of smoking’s apparent protective effect**

This study may afford a better understanding of the nature of cigarette smoking’s protective effect on PD risk. Therefore, it may help identify the specific pathways that influence smoking’s protective effect and help distinguish the chemical classes of neuroprotective or neurotoxic agents involved. If so, this may provide the potential to reveal new targets for strategies of altering PD risk.

- **Guidance for new strategies for studying other complex diseases**

In this project, we applied a statistically more powerful and more economic genetic epidemiological approach (case-only design) as yet unutilised in PD research. This approach circumvents many of the limitations of control ascertainment for the classical case-control study. Successful application of this approach may assist in opening a door, not only for PD, but also for research on other rare and late age-at-onset disorders. As is well known for these disorders, the time course and expense of recruiting subjects have made it extremely hard for a prospective study. Therefore, this project could provide guidance for new strategies to investigate the important interactions between genes and environment that constitute determinants of the most common complex diseases.
• **Better understanding of potential genetic determinants of smoking**

The major limitation of a case-only approach is the assumption that the exposure and the genetic factors examined are independent of one another. To date, no study has reported unquestioned correlation between cigarette smoking and the specific genotypes we are studying. In this thesis, we used a large group of unaffected subjects to examine the independence of the genotype-exposure relationship. Direct relationship between particular genotypes and cigarette smoking in unaffected individuals precludes the case-only approach to study the interactions with those genotypes. In this case, the data provide useful preliminary information that may inform future studies regarding genetic influences on smoking. Furthermore, any positive associations between genotype and smoking status would highlight limitations in previous case-control studies examining the association between these genetic factors and PD that failed to consider smoking as a potential confounding factor.
CHAPTER 2: LITERATURE REVIEW: CIGARETTE SMOKING, SPECIFIC GENETIC POLYMORPHISMS OF METABOLIC ENZYMES AND THE RISK OF PARKINSON’S DISEASE

2.1 INTRODUCTION

Parkinson’s disease (PD) is a major aging-related neurodegenerative disorder. At present, there is no specific cure for this disorder. Since the Australian population is aging, PD has become a significant public health concern in Australia. The aetiology of PD is still unclear. Both environmental and genetic factors may contribute to the development of PD. Currently it is suggested that this multifactorial disorder results from environmental factors acting on genetically susceptible individuals with normal aging as the risk modifier [71, 105-107]. Among environmental factors, cigarette smoking has been specifically of interest since epidemiological studies observed that smoking was negatively associated with the risk of PD.

The familial aggregation tendency has been observed in some PD patients. This observation triggered the search for genetic risk factors of the disease [108]. The study of rare cases of familial parkinsonism identified mutations in the alpha-synuclein gene and indicated an autosomal dominant mode of inheritance [64, 65]. Mutations in the parkin gene account for some cases of autosomal recessive early-onset disease [67, 69, 109]. Mutations of the Tau [70] and UCL-1 genes [66] also appear to be associated with some cases. However, the direct link between genetic mutations and familial PD account for a very small portion of PD cases [71, 72], while the majority of idiopathic PD still cannot be linked to mutations directly.
It is reasonable to consider that, if environmental factors such as chemical compounds found in cigarette smoke influence PD risk, differences in the ability of certain individuals to metabolise these substances might alter their susceptibility to developing PD. Clearly, it is very important to understand if there are any potential interactions between smoking and some specific genetic polymorphisms. The possibility exists that even if each factor, individually, might contribute only slight variation and influence a small portion of the whole population, the potential non-linear and unpredictable interactions may account for a high proportion of the aetiological fraction. Understanding this may therefore provide insights into the risks posed by potentially damaging endogenous or exogenous substances and help understand the unknown pathogenetic mechanisms. It may also lead to uncovering the environmental and genetic factors related to PD.

This chapter provides a systematic review of the literature related to the studies on cigarette smoking, specific metabolic enzyme genetic polymorphisms and their interactions in the aetiology of Parkinson’s disease (PD).

2.2 CIGARETTE SMOKING AND PD RISK

2.2.1 Epidemiological Studies

The effect of cigarette smoking on PD risk has long been of scientific interest since Dorn (1959) published the first study suggesting an inverse association [110]. Since then it has been consistently noticed that fewer PD patients have been smokers than controls [111-113]. A large number of epidemiological studies investigated the effect
of smoking on the development of PD. Early studies using the case-control design confirmed that PD patients were less likely to have smoked than controls and suggested a protective effect of smoking on PD [33, 52, 59, 114-117]. A meta-analysis also reported an inverse association between smoking and PD with an Odds Ratio (OR) around 0.5 (95%CI: 0.4-0.7) [60]. However, some researchers doubted the protective association and thought it could be an artefact. They argued that those early studies either did not adequately verify the diagnosis of PD, or did not adjust for other important confounding factors like gender and age in the analysis, or recruited control subjects inappropriately [118-122]. They also argued that none of these early studies could find a dose-response relationship, indicating that heavier smokers got greater protection. Indeed, theoretically, if cigarette smoking can exert a protective effect on PD risk with a relatively strong biological basis, a dose-response would be expected to exist, i.e., the incidence of PD in smokers, especially moderate and heavy smokers, should be significantly lower than lighter smokers, and smokers who did get PD should have an older average age at onset of PD and their clinical signs and symptoms should be less severe.

One example is the study by Mayeux and colleagues [123]. This study, using a case-control design, revealed that the ages at onset were quite similar for PD among both smokers and non-smokers; the odds ratios for a history of smoking associated with PD were 0.7-1.8 (refer to non-smokers), and no protective gradient was associated with heavier smoking patterns. They argued that the decreases in smoking were most likely a consequence of PD. However, in their study, there was a significant difference in gender, with more men among PD cases (50%) than among controls (19%), and there were also significant differences in age, education and ethnic
background between cases and controls. These factors were not well-adjusted in their analyses.

Many alternative explanations for the negative (protective) association between smoking and risk of PD have been suggested [60, 117, 119, 120, 124-127]. These include:

1. Smokers probably have a shorter life expectancy than non-smokers. Therefore, fewer PD cases are expected in smokers as the incidence and prevalence of PD increase dramatically with age, especially over 65;
2. In PD patients, mortality in smokers may be earlier than in nonsmokers because of a more rapidly fatal course in PD patients who continue to smoke. If this is true, due to the decreased survival among PD-smokers, they would be under-represented in a case-control design;
3. Individuals predestined to get PD tend to be more passive, more introspective, less likely to take risks, and more self-controlled, and in turn may choose not to smoke;
4. In the very early presymptomatic stage of their neurodegenerative conditions, the patients-to-be may have lost their interest in smoking;
5. In PD patients, there may exist a change in personality or a so-called ‘premorbid attitude’ that may be related to the cessation of smoking;
6. In PD patients, physical or mental limitations may result in the cessation of smoking;
7. Smoking may suppress the signs and symptoms of PD and lead to case-control misclassification;
8. Common lifestyles associated with smoking, such as coffee intake or tea drinking, have also been associated with lower incidence of PD;

9. Unmeasured confounding factors such as undiscovered genetic risk factors that predispose a person to acquire PD might simultaneously reduce the likelihood of smoking;

10. A recall bias may be operating, which is a methodological problem common to all case-control studies; and

11. Studies showing no inverse relationship between smoking and PD risk may not be published or may only be published in local journals or conference abstracts and proceedings.

To clarify this issue, a large number of studies, including several prospective cohort studies, have tried to provide further evidence against the above alternative explanations for the inverse association of smoking and PD risk. The prospective cohort study of 26-year follow-up by Grandinetti and colleagues [128], which has eliminated the potential for outcome-associated bias in the recall of smoking behaviours, revealed an apparent inverse dose-response effect on the risk of idiopathic PD through examination of smoking by pack-years. They also excluded the potential bias associated with long-standing health consciousness by comparing the frequencies of hospitalisation between smoking and non-smoking PD subjects, and excluded the potential bias by comparing other correlated behaviours of smoking like alcohol and coffee drinking. The limitation of this study is, as they pointed out, that persons who attribute parkinsonian symptoms to natural aging may never seek medical attention or receive a diagnosis. As Schoenberg and colleagues [46] reported, only 42% of
prevalent PD cases identified in a door-to-door survey could be identified via medical records and physician reporting.

The Rotterdam study recruited a population-based cohort with all the participants aged 55 years or older [129, 130]. This study screened 6969 participants for parkinsonian symptoms at both baseline and follow-up. The screened positives were further examined by a neurologist for PD diagnosis. Specifically, this study reported on analyses of smoking and incident idiopathic PD and yielded a relative risk estimate of 0.4 (95%CI: 0.2-0.9) for ever smokers [130].

Recently, Allam and colleagues systematically reviewed prospective studies that focused on examining the associations between cigarette smoking and PD. They included seven studies which were considered to be well designed and carefully executed. The pooled risk estimate for ever smokers was 0.5 (95%CI: 0.4-0.6), suggesting a significant inverse association [131].

Several carefully designed case-control studies also tried to test the above alternative explanations for the protective effect of smoking on PD risk. In order to minimise recall bias, Hellenbrand and colleagues [119] only recruited PD cases aged 65 years or younger. For each of the 380 cases, they recruited two age- and gender-matched controls, with the first one an immediate neighbour and the second one from the same urban or rural region. The data were collected by professionally trained interviewers with the cases and the corresponding controls being interviewed by the same person. Analysis of the data indicated that the age started cigarette smoking did not differ between cases and controls and the OR for ever smoked was 0.5, again conferring a
protective effect. Further analysis revealed that the inverse association between smoking and PD was still statistically significant even though only the amount smoked 20 years prior to PD diagnosis was considered. Similarly, a strong inverse relationship was found between pack-years smoked and PD when the analysis was stratified according to disease duration or age at onset. Their study also showed that adjustment for education, coffee or alcohol consumption had almost no effect on the inverse association.

Gorell and colleagues, using a population-based case-control study, examined the inverse association between cigarette smoking and PD among 144 cases and 464 age-, gender-, and race-matched controls. They detected an inverse dose-response relationship among ever-smokers and ex-smokers [57]. Further analysis minimized the likelihood that this inverse dose-response relationship was due to bias or confounding.

More recently, Galanaud and colleagues [132] further examined this question in a population with high prevalence of pesticide exposure. As is well known, pesticide exposure has been associated with the increase of PD risk [42, 45, 48]. Among the 247 cases and 676 age-, gender- and region-matched controls, a significant inverse association with smoking was detected, with an OR of 0.6 (95%CI: 0.4-0.9). The strength of this inverse association increased with the number of pack-years smoked. Moreover, including only smoking as long as 40 years before PD was diagnosed, the protective effect of cigarette smoking still could be detected.
A recent meta-analysis by Hernan and colleagues summarised the findings from 44 case-control and 4 cohort studies. They reported that the OR for ever-smokers was around 0.6 (95%CI: 0.5-0.6) whereas ORs for current- and past-smokers were 0.4 (95%CI: 0.3-0.5) and 0.8 (95%CI: 0.7-0.9), respectively [55].

The finding of smoking’s potential protective effect against PD has also been confirmed in populations other than Caucasians. Recently, a systemic review of studies in Chinese populations revealed a pooled OR of 0.8 (95%CI: 0.6-0.9) [133]. The authors suggested that the difference in the protective magnitude may result from racial differences, such as different genetic backgrounds between the populations.

In addition to empirical evidence from the studies, there are novel theoretical arguments against the alternative explanations. First, even if there is no dose-response gradient, it is likely that only major determinants of the clinical presentation of a disease will show such a risk gradient. In fact, even a major disease determinant may only have a threshold effect or a ceiling effect rather than a dose-response relationship. Besides, factors related to disease severity may or may not be the same as those that cause a disease to appear [55].

In summary, the inverse association between cigarette smoking and PD risk has been consistently reported from studies with different epidemiological designs with a significant inverse dose-response relationship observed both for the duration of smoking and for the total pack-years smoked. In addition, the inverse association is apparent for smoking that occurred long before PD was diagnosed and appears not to
be explained by confounding factors. Therefore, it is reasonable to believe that cigarette smoking confers a true biological protective effect on the development of PD.

2.2.2 Potential Mechanism of the Protection

Processed tobacco contains over 3000 compounds. After a cigarette burns, the mainstream and sidestream tobacco smoke might contain more than 4000 constituents, with the major ones being carbon monoxide, nicotine, and tar [134]. A number of experimental studies have investigated the potential mechanism underlying the inverse association with PD and suggested that at least three possibilities should be considered:

1. By a chemical and/or biochemical process, substances in cigarette smoke may exert a direct or indirect protective effect on nigral neurons or glia to reduce the damage to them;
2. By altering the metabolic enzyme activities and/or by competing with original endogenous substrates of enzymes, substances in cigarette smoke may alter the metabolism of endogenous neurochemical compounds; and
3. By altering the metabolic enzyme activities and/or by competing with other exogenous substrates of enzymes in the brain and/or other organs, substances in cigarette smoke may inactivate otherwise neurotoxic substances.

The majority of studies were focused on nicotine and found it may play a role in the protective effect on PD risk. Nicotine is absorbed through the lung epithelium and travels directly to the brain, where it readily crosses the blood-brain barrier. Ishikawa
and colleagues [135] reported that cigarette smoking and nicotine gum chewing in persons who already smoked could transiently relieve tremor, rigidity and bradykinesia. In animals, the administration of nicotine may exert protection against the cell loss of nigral dopaminergic neurons induced by partial mesodiencephalic hemitransection [136]; chronic nicotine treatment may also reduce the cell loss of age-related nigrostriatal neurons [137]. Nicotine has been noticed to stimulate the release of dopamine and it might also directly interfere with the uptake of neurotoxic compounds like MPP+ ions [138-140]. It is possible that nicotine’s neuroprotective effect may come from its activation of alpha-4 nicotine receptors [141]. In addition, Shahi and colleagues [101] noticed that nicotine’s protection against MPTP-induced striatal dopamine depletion in mice was, to some extent, due to nicotine-inducing cytochrome P450 enzymes. They suggested that this induction might accelerate the metabolism and detoxification of possible neurotoxins such as MPTP, and thus indirectly confer the protective effect.

Smoking could reduce human brain MAO-B activity substantially, as assessed by positron emission tomography (PET) [142]. Since MAO-B is involved in the biotransformation of MPTP-like endogenous or exogenous compounds to the active neurotoxins (eg, MPTP is converted to MPP+), the potential exists that the inhibition of this enzyme may change an individual’s response to MPTP-like neurotoxins [143, 144]. Moreover, MAO-B is responsible for oxidisation of dopamine and this process could lead to the formation of hydrogen peroxide [145]. As oxidative stress in critical areas of the brain has been suggested to play an important role in the pathogenesis of PD [4, 5], the inhibition of MAO-B may potentially reduce the oxidative stress in the brain, and thus confer protection against PD.
Besides nicotine, another common component of cigarette smoke is carbon monoxide. It has been suggested that carbon monoxide in cigarette smoke can act as a free radical scavenger [17]. Sagone and colleagues found that carbon monoxide appears to protect red blood cells from hydrogen peroxide-induced membrane damage [146]. If this also happens in substantia nigra, carbon monoxide generated by cigarette smoke would create a specific mini-environment that might help protect nigral cells from the attack of free radicals.

Studies have also indicated that hydrazine and 4-phenylpyrodine, components of cigarette smoke, provide partial protection of dopaminergic nigrostriatal neurons in mice given MPTP [144, 147]. The authors suggested that the protection could come from a competitive mechanism. In addition, Dowskin and colleagues noticed that other quaternary N-methylated nicotine derivatives may also present a similar effect [148]. It is reasonable to postulate that the accumulation of these protective substances in dopaminergic nigrostriatal neurons of cigarette smokers might reduce smokers’ susceptibility to MPTP-like neurotoxins and thus reduce the likelihood of developing PD.

2.3 GENETIC INFLUENCE ON CIGARETTE SMOKING

Smoking behaviour can be modulated by genetic factors. Carmelli and colleagues conducted a twin study [149] which compared the concordance for smoking between monozygotic and dizygotic twins. The authors assessed the relative contribution of familial and genetic factors. The ratio of observed to expected concordance for
smoking was higher among monozygotic twins than among dizygotic twins for those who had never smoked (RR: 1.4; 95%CI: 1.3-1.5), for former smokers (RR: 1.6; 95%CI: 1.4-1.9), for current cigarette smokers (RR: 1.2; 95%CI: 1.1-1.3), and for current cigar or pipe smokers (RR: 1.6; 95%CI: 1.2-2.1). Their data also suggested genetic influences on quitting smoking. Monozygotic twins were more likely than dizygotic twins to be concordant for quitting smoking (OR: 1.2; 95%CI: 1.1-1.5). They concluded that in adult male twins, there were moderate genetic influences on lifetime smoking practices. Later on, True et al analysed data collected from 2,204 male monozygotic and 1,793 male dizygotic twin pairs and also confirmed the finding of a major genetic influence on smoking persistence [150].

Edwards et al examined the genetic influences on smoking behaviour in female twins and suggested that, similar to males, genetic traits might influence smoking initiation (RR: 1.6; 95%CI: 1.0-2.3) and maintenance (RR: 1.8; 95%CI: 1.3-2.4). After adjusting for age, education, and frequency of contact, the adjusted logistic regression models confirmed the findings [151]. More recently, in a study of 991 randomly selected current smokers, Niu and colleagues showed an OR of 2.1 (95%CI: 1.0-4.4) for nicotine-dependence in second siblings in families in which the first sibling was nicotine-dependent [152].

Some authors have attempted to uncover the underlying mechanism for smoking persistence. One of the examples is the double-blind study conducted by Sherwood et al [153]. This study investigated the psychomotor effects of cigarette smoking on simulation of driving comprising continuous tracking and brake reaction time tasks. The results indicated that cigarette smoking could decrease irritability, heighten
attention and improve driving performance. The authors suggested that an optimal nicotine dose may exist for the enhancement of cognitive and psychomotor function.

L-nicotine, which is the most pharmacologically active form of nicotine and is also the major constituent of tobacco, has been thought to be responsible for tobacco dependence. Individuals smoke cigarettes habitually to maintain their blood nicotine levels. When smoking is stopped, smokers may suffer from withdrawal syndromes, which include weariness, nausea, constipation, diarrhea, insomnia, and depressed concentration and psychomotor activity. These symptoms of withdrawal are improved by administration of nicotine.

Several studies have attempted to localise genes that contribute to smoking-related phenotypes. For this purpose, genome-wide linkage analysis is considered to be an important tool. Linkage findings can support the candidacy of known genes and hint at the identities of unknown phenotype-related genes. Three linkage analyses of smoking behaviour have been reported. Two used COGA (Collaborative Study on the Genetics of Alcoholism) families and one carried out a sib-pair linkage analysis. They reported some evidence for linkage of smoking behaviour to chromosome 15q, as well as GATA193 on chromosome 17p [154-156].

Straub et al examined families from New Zealand and Richmond, Virginia, in the USA. Regions on chromosomes 2, 4, 10, 16, 17, and 18 showed evidence of linkage to nicotine dependence. Among them, D2S1326 showed the highest lod score in the New Zealand families. These linkage findings were not replicated in the Richmond families. These data indicate that the effects of genes on smoking behaviour are weak,
or that genetic alleles that influence smoking behaviour occur in only a small proportion of families [157].

Genes involved in metabolism of nicotine are considered important to smoking-related behaviours. Nicotine is metabolised in three different ways: by C-oxidation, N-oxidation, and N-methylation, and there are important and interesting inter-individual differences in nicotine metabolism among humans. It is postulated that genes that might possibly be associated with smoking may include CYPs, GSTs, NATs, MAOs, dopamine receptors (D1, D2, D3, D4 and D5 receptors), dopamine transporter, serotonin transporter, serotonin receptors, tyrosine hydroxylase, tryptophan 2,3-dioxygenase, opioid receptors, and cannabinoid receptors.

2.4 HUMAN GENETIC POLYMORPHISMS OF METABOLIC ENZYMES AND PD RISK

Scientific interest in the effect of human genetic polymorphisms of metabolic enzymes on PD risk has been driven mainly by three observations: first, sequencing of metabolic enzyme genes revealed that many exhibit polymorphic traits and some of the variants might alter gene function. Second, the development of simple assays based on PCR and RFLP methods allowed easy and fast identification of individual genotypes for a variety of polymorphisms. Third, it is assumed that, if environmental factors, especially chemical compounds, are associated with PD risk, the difference in the ability of certain individuals to metabolise these substances might alter their susceptibility to developing PD.
2.4.1 Mechanisms of Genetic Polymorphisms Influencing PD Risk

Theoretically, polymorphisms in human metabolic enzyme genes may modulate gene function by a variety of molecular mechanisms, which include:

1. Nucleotide variations in the coding region of the gene resulting in amino acid substitution and altered enzyme activity or substrates;
2. Deletions in the coding region of the gene leading to an inactive enzyme or reduced protein synthesis;
3. Polymorphisms in the non-coding region affecting transcriptional control elements involved in basal enzyme expression and/or induction in the coding region;
4. Variations in the signalling region of the gene affecting transcript half-life and thus the quantity of enzyme;
5. Gene amplification increasing the quantity of enzyme; and
6. Complex interactions between various polymorphic genes or their various enzyme catalysis products.

There are a number of candidate genetic polymorphisms that fulfil one or more of these criteria and potentially interact with chemicals found in cigarette smoke or otherwise influence risk of PD.

2.4.2 Molecular Epidemiological Studies

Molecular epidemiological studies aimed at examining the effect of human genetic polymorphisms of metabolic enzymes on PD risk have shown that the presence of
some polymorphisms might be statistically significantly associated with PD risk [158]. The important concept is, although these associations do not necessarily imply direct causal relationships between the genetic factors and PD, the presence of statistically significant associations (if consistently reported) might be helpful for guiding further studies to elucidate the pathogenesis of PD.

2.4.2.1 Glutathione-S-Transferases (GSTs):

GSTs are a superfamily of enzymes involved in the detoxification of chemical compounds and metabolites. To date, six of the human cytosolic GST genes have been confirmed to be polymorphic, i.e., GSTM1, GSTM3, GSTO1, GSTP1, GSTT1 and GSTZ1. In human beings, different classes of GSTs may have different substrate specificities and tissue distributions [159], although to some extent, they may also overlap with each other. As glutathione-dependent enzymes, GSTs have direct antioxidant activity and are responsible for controlling the reduction/oxidation status of a cell. Theoretically, they might affect a variety of signal transduction pathways involved in the regulation of cell proliferation and apoptosis. In fact, experimental studies by Adler and colleagues have shown that GSTP1 interferes with this complex network [160, 161]. GSTs are also involved in the metabolism of dopamine [162-164]. The GST enzyme activity in the brains of people with PD has been reported to be similar to that in the brains of people without PD. However, it is possible that some specific classes of GSTs might distribute and express differently. If this is true, when people are exposed to specific environmental neurotoxicants, the reduction/oxidation status of cells may differ between individuals. In the following sections, I extensively review the literature related to association studies between GSTM1, GSTP1, GSTT1
and GSTZ1 and PD risk. GSTM3 and GSTO1 are not included since few published studies were available.

**Glutathione-S-Transferase M1 (GSTM1):**

GSTM1 is an inducible phase 2 enzyme. The gene product of GSTM1 is involved in scavenging many electrophilic reactive intermediates. Aminochrome-dependent neurodegeneration has been proposed as a mechanism of pathogenesis of PD. Excessive amounts of neurotoxic catecholamine-derived quinones such as dopachrome, adrenochrome and noradrenochrome, and the oxidative stress that accompanies their metabolism, might lead to PD [165-167]. It has been reported that GSTM1 catalyses the conjugation reaction between glutathione and catecholamine-o-quinones under physiological conditions [163-165, 168]. Thus, a defect in this GSTM1 might influence an individual’s predisposition to PD.

The human GSTM1 gene is located on chromosome 1p13. Individuals with one or two copies of the GSTM1 allele are designated GSTM1 positive (non deletion); individuals with homozygous deletion of the GSTM1 allele are designated GSTM1 negative (null or deletion). In humans, homozygous deletion of GSTM1 locus has been reported to affect about 40% to 60% of Caucasians as well as Asians [169, 170]. Theoretically the deletion can cause the loss of gene function and affect substantially the metabolism of some substances. A number of studies have tried to establish a link between the polymorphic expression of GSTM1 and the incidence of PD.
Ahmadi and colleagues [73] reported a significant, elevated median age for the onset of PD among GSTM1 gene carriers (median age = 68 years) compared to PD patients with GSTM1 null genotypes (median age = 57 years). The results indicated that carriers of the GSTM1 gene postpone the onset of PD, which suggests that the detoxication pathways might represent important protective mechanisms against reactive intermediates modifying the susceptibility and onset of PD. Particularly, Stroombergen and colleagues [171] reported that males with a deletion of the GSTM1 gene were more susceptible to PD, which implies that environmental factors which specifically target men may be involved.

In contrast, the study conducted by Rahbar’s research team did not replicate the correlation between GSTM1 deletion and age at onset of PD and did not support the hypothesis of a possible impact of GSTM1 detoxification activity in the pathogenesis of PD [172]. The advantage of this study is that the authors subdivided PD patients into two groups according to age at onset. Therefore, they excluded the possibility that risk levels may differ between subgroups.

Bandmann and colleagues did not detect an association between GSTM1 deletion and PD risk (for Idiopathic PD: OR=1.09, 95%CI: 0.61-1.93; for familial PD: OR=0.89, 95%CI: 0.50-1.57). This study took advantage of using only DNA samples extracted from brain tissue with pathological diagnosis of PD for the idiopathic PD group [173]. Therefore, their results are less likely to be biased by misclassification.

Recently, Harada and colleagues grouped GSTM1 into three genotypes (deletion / deletion, deletion / nondeletion and nondeletion / nondeletion), which might enable
examination of the effect of deletion more accurately. They found no association with PD risk [174].

To date, studies are still unable to provide sufficient information to judge if there is a link between polymorphic expression of GSTM1 and the risk of PD.

**Glutathione-S-Transferase P1 (GSTP1):**

The GSTP1 locus is located on chromosome 11q13. The polymorphisms at the GSTP1 locus were first reported by Board’s research team and result in amino acid substitutions [175, 176]. Two main GSTP1 gene variants were investigated: the A/G transition at nucleotide 313 which results in an isoleucine at position 105 being replaced by valine, and the C/T transition at nucleotide 341 leading to an alanine at position 114 being replaced by valine. As usually defined, GSTP1*A has isoleucine at position 105 and alanine at position 114, GSTP1*B has valine at position 105 and alanine at position 114, and GSTP1*C has valine at both positions 105 and 114. The gene variations of GSTP1 have been reported to affect substrate selectivity and stability [177-179], and result in defective enzyme activity [180-183]. In human beings, GSTP1 is present in the brain and blood-brain barrier [184].

Menegon and colleagues first investigated the possible connection between GSTP1 polymorphism and PD [81]. They reported that the A/B/C frequencies were 0.63:0.31:0.06 in PD patients and 0.65:0.28:0.06 in controls. Therefore, no association was found between GSTP1 polymorphisms and idiopathic PD. However, there were significantly more PD patients who had GSTP1 gene variants among those with
pesticide exposure. Among these people, GSTP1*A was protective. The authors suggested that polymorphisms alter the ability of the enzyme to detoxify pesticides that may have neurotoxic effects [81]. However, one of the drawbacks of this study is the small sample size of 95 PD cases and 96 controls.

More recently, Kelada and colleagues [97] further examined the potential association between GSTP1 polymorphism and PD. They recruited 214 idiopathic PD cases and 330 age- and gender-matched controls. Interestingly, they found that distributions of GSTP1 genotype between cases and controls were slightly different. There were more GSTP1 Ile104Val heterozygotes among cases than among controls (OR: 1.43, 95%CI: 0.98-2.08). They suggested that GSTP1 polymorphisms might play a minor role in idiopathic PD.

Summarily, studies to date suggest that there might be a potential association between GSTP1 polymorphisms and the risk of PD. However, it is also possible that association between GSTP1 gene variants and PD susceptibility may be secondary to linkage disequilibrium with other genes near the GSTP1 locus, such as aldehyde dehydrogenase. Clearly, further studies are needed to clarify this point.

**Glutathione-S-Transferase T1 (GSTT1):**

Chemical compounds in tobacco and their metabolites, reactive oxygen species and lipid peroxidation products, are likely to be substrates for GSTT1 detoxifying enzymes. GSTT1 is an inducible enzyme. In humans, GSTT1 gene is located on Chromosome 22q11. Homozygous deletion of the GSTT1 locus has been reported to
affect about 15% to 25% of Caucasians as well as Asians [169, 170]. The homozygous deletion might cause the loss of gene function and affect substantially the metabolism of some substances. Interestingly, GSTT1 has also been known to activate certain industrial chemicals, such as dihalomethanes and 1,2-dibromoethane [185].

Palma and colleagues reported that the GSTT1 null genotype had a significant, positive association with PD risk (OR: 1.8, 95%CI: 1.03-3.15) [186], whereas Bandmann’s research team reported no association (for idiopathic PD, OR: 1.07, 95%CI: 0.51-2.26; for familial PD, OR: 1.30, 95%CI: 0.63-2.68) [173]. As mentioned previously, the study conducted by Bandmann and colleagues took the advantage of using only DNA samples extracted from brain tissue with pathological diagnosis of PD for the idiopathic PD group. Similarly, a study conducted by Menegon et al [81] and a more recent study conducted by Kelada et al [97] did not detect any association between GSTT1 deletion and PD either. In addition, the study conducted by Rahbar and colleagues also reported no correlation between GSTT1 deletion and the age of onset of PD [172].

Hence, currently there is no evidence to support the hypothesis of a possible impact of GSTT1 detoxification activity in the pathogenesis of PD.

**Glutathione-S-Transferase Z1 (GSTZ1):**

The GSTZ1 locus is located on chromosome 14q24. The polymorphisms at the GSTZ1 locus may result in amino acid substitutions. Mainly two GSTZ1 gene
variants have been investigated: an A/G transition causes Lys at position 32 to be replaced by Glu, and another A/G transition causes Arg at position 42 to be replaced by Gly.

To date, only one study by Menegon and colleagues has been conducted to investigate the role of GSTZ1 in PD susceptibility [81]. The A/G ratio for Lys32Glu was 1:2.39 in PD patients and 1:2.52 in controls while the A/G ratio for Arg42Gly was 1:18 and 1:13.6, respectively. Thus, they did not find an association between these two GSTZ1 polymorphisms and PD risk.

**Overall Summary:** Results of the studies examining the association between GST polymorphisms and PD risk are presented at the end of this chapter (Table 2.1.). So far, the results have been conflicting. One of the problems might come from different GST isoenzymes being known to exhibit overlapping substrate specification.

### 2.4.2.2 Cytochrome P450s (CYPs)

Cytochromes P450s (CYPs) are a superfamily of heme-containing enzymes. As the most versatile and important phase 1 metabolic enzymes, CYPs are involved in the oxidative metabolism of a variety of endogenous and exogenous lipophilic compounds. According to their amino acid sequence similarities, CYPs are grouped in subfamilies. CYPs are named with the root CYP followed by a number designating the family, a capital letter that belongs to the subfamily and a number for the individual form. For example, CYP2D6 denotes family "2", subfamily "D" and individual member "6". To date, 57 functional CYP genes and 46 pseudogenes
CYPs are the target of special interest in PD studies because most, if not all, neurotoxic chemical compounds are either substrates, inducers, or inhibitors of this enzyme family. CYPs catalyse the oxidative biotransformation of lipophilic substrates to less hydrophilic metabolites, and the products further undergo conjugation reactions with glucuronide, sulfate, glutathione, etc., in phase 2 metabolism. The product of this reaction is generally hydrophilic and is readily excreted from the body. No enzyme of this family has been found to possess unique substrate specificity and many may perform more than one catalytic function on more than one substrate. Similarly, multiple CYP enzymes may act on one chemical or substrate. However, substrate selectivity and stability could be different for individual CYPs. The following sections only review the literature related to association studies between CYP2D6 or CYP2E1 and PD risk, since few published studies were available for other CYPs.

**Cytochrome P450 2D6 (CYP2D6):**

The human CYP2D6 gene is located on chromosome 22q13. As a phase one enzyme, CYP2D6 is involved in the biotransformation of various kinds of chemical compounds, including oxidising tetrahydroisoquinolines to neurotoxic isoquinolinium ions, and metabolising exogenous PD-inducible neurotoxin MPTP. A study by Gilham and colleagues has demonstrated that MPTP N-demethylation, a detoxification pathway, is catalysed by CYP2D6 and up to 40% of the hepatic metabolism is mediated by this enzyme [187]. The CYP2D6 enzyme is expressed in the human brain. *In situ* hybridization has demonstrated that CYP2D6 is also
localized in the pigmented neurons of the substantia nigra, indicating that 2D6-mediated detoxification may occur in target cells [187]. It is speculated that the abnormality of CYP2D6 activity may cause weak detoxification or activation of some environmental neurotoxins with similar structures and toxicity to MPTP, and that, as a result, it might cause accumulation in the brain. Therefore, CYP2D6 can be a factor in susceptibility to MPTP neuronal toxicity, and altered gene function in CYP2D6 may be associated with PD [187].

CYP2D6 is a non-inducible enzyme. Lessened gene function is mainly caused by two point mutations that result in CYP2D6*3 (formerly A) and CYP2D6*4 (formerly B) alleles while loss of gene function is mainly caused by deletion of the entire CYP2D6*5 allele (formerly D) [188]. CYP2D6*3, CYP2D6*4 and CYP2D6*5 may account for more than 90% of the poor metabolisers in Caucasians, with an autosomal recessive mode of inheritance.

CYP2D6 is also involved in activating nicotine and tobacco-specific nitrosamine 4-methyl nitrosamino-1.3-pyridyl-1-butanone, and in causing oxidative stress through the generation of free radicals. Extensive metabolisers are likely to generate more free radicals. Furthermore, the majority of CYP2D6 substrates are small molecules that interact with the protein via an electrostatic interaction between basic nitrogen that is common to the majority of CYP2D6 substrates and an aspartic acid residue in the active site of the protein. Therefore, it is reasonable to speculate that the alteration of CYP2D6 gene function may play a role in PD.
Allele frequencies of CYP2D6 have been shown to vary widely between populations of different racial origins. Joost and colleagues reported that the frequencies of the CYP2D6*3, CYP2D6*4 and CYP2D6*5 mutant alleles in Caucasians were 0.9%, 20.5% and 2.7% [78], while Chida et al reported that the frequencies of the mutant alleles were 0%, 0.8% and 4.1% in Japanese populations [189]. Based on their data, Chida and colleagues estimated that the population frequency of the CYP2D6 poor metaboliser (PM) phenotype was 0.29% in Japanese people. A Korean study [89] and several studies in Chinese populations [80, 190, 191] reported similar results to the Japanese populations. CYP2D6 poor metabolisers are estimated to represent around 8% of Caucasians [78, 82]. Therefore, compared to Caucasian populations, the poor metaboliser genotype is extremely rare in Asians.

Armstrong and colleagues first reported that the most common mutant allele, CYP2D6*4, was twice as frequent among patients as in controls, with an approximate relative risk of 2.7 (95% CI: 1.1-6.4) for subjects homozygous or heterozygous for this allele [192]. Similarly, Kurth and colleagues reported an OR of 1.9 (95% CI: 1.3-2.4) [193] while Lucotte et al, using region-, sex- and age-matched control subjects, reported an OR of 2.1 (95%CI: 1.4-2.6) [194]. Atkinson and colleagues genotyped a large series of clinically and neuro-pathologically confirmed PD cases and age-matched control individuals for the CYP2D6*4 allele and also reported an elevated frequency in PD cases [74]. Meanwhile, Akhmedova et al found that only a particular subset of PD patients with akinetic-rigidity tremor symptoms had increased frequency of CYP2D6*4 and suggested that CYP2D6 polymorphism might predispose the carriers to a specific form of PD [195]. A study showed that even though the heterozygous genotype did not increase the risk of PD, those with the homozygous
mutations had a 2.4- to 2.5-fold increase in PD risk. McCann and colleagues reported that CYP2D6 poor metabolisers were more common among PD patients than among matched controls (OR: 1.7; 95% CI: 0.9-2.5) [196]. A meta-analysis of these results by McCann et al, together with ten other published studies, gave a pooled odds ratio for CYP2D6 PM of 1.5 (95% CI: 1.2-2.0; p=0.01). They suggested that CYP2D6 PM might have a small but highly significant association with PD. However the association could be easily missed in small studies [196].

There were also some authors who questioned the link between CYP2D6*4 or poor metaboliser status with PD risk. Christensen et al [75] and Rostami-Hodjegan et al [83] have separately re-evaluated the primary research on the association between the CYP2D6 polymorphism and the risk of PD using meta-analysis. The OR calculated by Christensen was 1.5 (95% CI: 1.1-2.0) and the analysis by Rostami-Hodjegan gave an OR of 1.3 (95% CI: 1.0-1.8), which was of borderline statistical significance (p < 0.074) and similar to the report by McCann et al; however, both authors indicated that if a single large study included in their meta-analysis that was statistically significant was excluded, the p-values would increase greatly to be non-significant. It is possible that the same is true of McCann’s meta-analysis. Therefore, Christensen and Rostami-Hodjegan conclude that there is no convincing evidence of an association between the CYP2D6 polymorphism and PD. Both authors also suggested that additional well-designed studies may allow a definitive conclusion, although any risk associated with CYP2D6 PM is likely to be small and therefore of questionable clinical significance. An important lesson from their analyses is that much time, effort, expense and patient inconvenience might have been avoided if more attention had been paid to appropriate study design particularly in the selection of control groups.
More recently, Maraganore et al investigated a sample of 139 unrelated PD cases and 113 control subjects. They did not detect an association between CYP2D6 PM and PD risk [197]. Similarly, Harhangi’s research team designed a case-control study nested within a prospective population-based cohort study in which cases and controls were sampled from the same source population. They investigated the two mutant alleles of CYP2D6*3 and CYP2D6*4 and the wild-type allele in 80 PD patients and 156 age- and gender-matched controls. Their results were not consistent with the hypothesis that the CYP2D6 gene is a major gene responsible for PD [77].

Lucotte and colleagues reported that frequency of the CYP2D6 mutations was slightly enhanced after age-at-onset of 60 years [194]. A study by Payami et al reported that mean onset age was significantly later in CYP2D6*4-positive patients while the frequency of CYP2D6*4 was significantly higher in late-onset PD than early-onset PD [82]. When they analysed the early- and late-onset PD separately, CYP2D6*4 had no effect on onset age; hence, they suggested that the association with delayed onset was more likely to be an artefact of an elevated CYP2D6*4 frequency in late-onset PD. Their results, contrary to a common assumption that CYP2D6 frequencies did not change with age, revealed that CYP2D6*4 frequency rose significantly with advancing age, both in patients with PD and subjects without PD. They also found that CYP2D6*4 frequencies in patients with early- and late-onset PD were in agreement with similarly aged subjects without PD, although different from each other, suggesting the elevated CYP2D6*4 frequency in late-onset PD was likely an age effect unrelated to PD. They concluded that the CYP2D6*4 allele is not associated with earlier PD onset. Instead, CYP2D6*4 might be associated with
survival. They suggested that inconsistent results from allelic association studies might have been due to an unrecognised age effect.

Wilhelmsen and colleagues examined the association between CYP2D6 polymorphisms and PD in a case-control study that included 10 polymorphic dinucleotide repeat markers linked to CYP2D6 to determine whether the association was present or due to linkage disequilibrium [198]. They found no association between any polymorphism of CYP2D6 and PD, but two of 10 dinucleotide repeat markers linked to CYP2D6 were associated with the disease. Their results suggested that there might be an unidentified locus for susceptibility to PD that is in linkage disequilibrium with dinucleotide repeat markers mapping near CYP2D6 on chromosome 22q13.

Stefanovic et al took the advantage of dividing PD patients into two subgroups according to Hoehn and Yahr staging (HY), a scale for evaluation of the severity of PD. They observed that the frequency of CYP2D6*4 allele was higher in PD patients than in a control group (OR=2.1, 95%CI: 1.1-4.0). Moreover, in the HY >2.5 subgroup, ie, patients with more severe symptoms and disabilities, the CYP2D6*4 allelic difference was even greater (OR=2.7, 95%CI: 1.1-6.6) [86], suggesting that CYP2D6*4 allele might be associated with the severity of PD.

Bon and colleagues specifically focused on patients with morphologically confirmed PD. They obtained the brain tissue from 50 patients. They also collected blood samples from 149 patients with clinical parkinsonism and from 96 healthy control subjects. A significantly higher allele frequency of CYP2D6*4 was found in patients
with PD (35%) but not with non-idiopathic parkinsonism (14.1%), compared to control subjects (19.8%) [199].

Results of the studies examining the association between CYP2D6 polymorphism and PD risk are presented at the end of this chapter (Table 2.2.). On the whole, critical analyses of the literature and recent studies emerging from independent laboratories have not confirmed that CYP2D6 polymorphisms confer susceptibility to PD risk. Conflicting results among reported studies of the prevalence of mutations among patients with PD suggest it is important to carry out more comprehensive studies and to analyse the interactions between this gene and other suspected risk factors.

Cytochrome P450 2E1 (CYP2E1):

The human CYP2E1 gene is located on chromosome 10q24. The CYP2E1 enzyme is expressed in brain and induced by structurally unrelated chemicals, such as ethanol and benzene [200-202]. More than 70 different substrates are specifically metabolised by CYP2E1 enzyme. Among them are most organic solvents, paracetamol and several precarcinogens such as aniline, benzene, dimethyl nitrosamine, vinyl chloride and urethane [203-206].

CYP2E1 may cause oxidative stress, and the oxy radicals generated by this enzyme are able to initiate NADPH-dependent lipid peroxidation. CYP2E1 metabolises N-hexane, leading to the formation of its neurotoxic metabolite 2,5-hexanedione. It has also been shown that enhanced activity of CYP2E1 might form isoquinolines from acetaldehyde (a product of ethanol metabolism) and dopamine [207-209]. Thus, any
functional polymorphism of this enzyme might be an important factor associated with the predisposition to chemically-induced PD.

It has been demonstrated that there are marked individual differences in the expression of CYP2E1 in human beings at the immunochemical level and to a lesser extent in catalytic activity assessed [210, 211]. Analysis and characterization of polymorphisms in the CYP2E1 gene have confirmed polymorphic sites at –1019 bp in the 5’-upstream region (Rsa I or c1/c2 allele) that are generally linked and also associated with transcriptional regulation of this gene [212], resulting in different expression levels of the CYP2E1 mRNA. Hayashi and colleagues (1991) reported that the RsaI allele is associated with a 10-fold higher concentration of CYP2E1 transcription [212]. A study by Montoliu et al reported that induction of the CYP2E1 enzyme might cause oxidative stress in astrocytes, leading to increased concentrations of malonaldehyde (a late-stage marker for lipid oxidation) and decreased concentrations of glutathione [213]. Similar changes have been seen in PD. Most interestingly, Riedl and colleagues found that CYP2E1 was specifically localised in the substantia nigra [214].

Similar to CYP2D6, allele frequencies of CYP2E1 have also been reported to vary widely between populations of different racial origins. The frequencies of the RsaI c1/c2 were reported to be from 4% (in Caucasians) to 20% (in Asians) [212, 215-217].

Few molecular epidemiological studies have been carried out to investigate the link between CYP2E1 RsaI polymorphisms and PD risk. A study conducted by Bandmann and colleagues obtained the ORs of 1.7 (95%CI: 0.6-4.9) for idiopathic PD and 1.5
(95%CI: 0.5-4.5) for familial PD, respectively [173]. The advantage of this study, as mentioned previously, was that they only included DNA samples extracted from brain tissue with pathological diagnosis of idiopathic PD. Therefore, it may have avoided the pitfall that 20% of patients with the clinical diagnosis of PD might have a different underlying pathology at necropancy [218, 219]. However, due to the small sample size, the authors could not identify if there is a significant association between CYP2E1 Rsal polymorphism and PD risk. Another study, conducted by Wang and colleagues, also reported that the Rsal polymorphism was not associated with PD risk in a Chinese population, even in patients with early-onset PD (onset <= 50 years of age) in which genetic factors might be more important [220].

Two CYP2E1 gene variants were also found with functional mutations in exon 2 and exon 8. The first one is a G1168A point mutation which causes an R67H amino acid substitution and the second one is a G10509A base substitution which yields a V389I amino acid exchange [221]. These variant alleles are very rare and probably not of particular public health significance.

An overall summary of the studies examining the association between CYP2E1 Rsal polymorphism and PD risk are presented at the end this chapter (Table 2.3.). To date, there are no sufficient data to judge if there is an association between the polymorphic expression of CYP2E1 and the risk of PD.
2.4.3 Combinations of Genetic Polymorphisms

The human metabolic enzyme system has been proven to be very complex. It is possible that the genetic polymorphisms of different metabolic enzyme genes may play different roles in the development of PD when in different combinations. Thus, it is important to consider the effect of genotypes across loci since evaluation of single loci may not be sufficient given the complementary and redundant nature characterising the human metabolic system. However, this is very difficult due to the requirement of extremely large sample sizes and also due to the potential complexity of the gene-gene interaction itself.

A study conducted by De Palma et al noted that although CYP2D6 polymorphisms alone had no significant association with PD risk, there was a significant association for the combined haplotype of CYP2D6 extensive metaboliser/GSTT1 null genotype (OR: 2.4, 95%CI: 1.3-4.6) [186]. The underlying mechanism could be due to CYP2D6 extensive metabolisers being likely to generate more free radicals, and in the central nervous system, these free radicals are generally detoxified by conjugation with glutathione, a reaction catalysed by GSTs.

Similarly, in a series of unrelated idiopathic PD patients, Santt and colleagues [222] observed a significant multiplicative effect for CYP2D6 poor metabolisers with GSTM1 null genotype. This study detected a 9.5-fold increase of PD risk among patients with the haplotype of CYP2D6 poor metabolisers/GSTM1 null genotype, compared to patients of non-CYP2D6 poor metabolisers/non-GSTM1 null genotype. The authors suggested that patients of CYP2D6 poor metabolisers/GSTM1 null
genotype may be less effective at detoxifying neurotoxins such as MPTP, and thus, more susceptible to PD.

2.5 GENETIC POLYMORPHISMS OF METABOLIC ENZYMES AND CIGARETTE SMOKING IN COMBINATION AS A RISK MODIFIER OF PD

It is reasonable to postulate that if smoking status significantly alters PD risk, it is possible that the previous inconsistencies in genetic association studies might be, at least partially, caused by this confounding factor as none of the previous association studies have adequately considered smoking status. Meanwhile, very few studies have tried to examine the hypothesis that polymorphisms in metabolic enzyme genes might modulate the effect of smoking on PD risk. Interestingly, the few papers that considered potential interactions between smoking and family history or genetic polymorphisms and risk of PD have indicated that genetic factors may significantly modify the inverse association with smoking, either increasing or decreasing the protection.

De Palma and colleagues reported that the protective effect of tobacco smoking (OR: 0.5, 95%CI: 0.3-0.9) was lost among patients with homozygous GSTM1 deletion (OR: 1.0, 95%CI: 0.5-2.0) while smoking showed a stronger inverse association with PD among patients with the GSTM1-positive genotype (OR: 0.2, 95%CI: 0.1-0.6) [186]. The authors suggested that the interaction might rely on the inability of GSTM1-null patients to detoxify free radicals formed from metabolites of substances associated with nicotine or from the dopaminergic activation induced by nicotine, which induces the production of free radicals.
More recently, in an age- and gender-matched case-control study, Kelada and colleagues noticed that although there was no significant association between GSTP1 genotypes and PD risk in general, the stratified analysis detected a significantly different distribution of GSTP1 Ile104Val heterozygotes among cases who ever smoked (OR: 1.92, 95%CI: 1.12-3.29) compared to cases who never smoked (OR: 1.04, 95%CI: 0.60-1.83) [97]. There are at least two advantages to this study: first, they recruited a relatively large sample of 214 cases and 330 controls; second, all cases were newly diagnosed with idiopathic PD.

To date, the number of studies on the interaction between smoking and genetic polymorphisms is still far from sufficient to yield a conclusion regarding an association with PD risk. However, the above studies have indicated that it is important to consider such potential interactions when we try to elucidate the aetiology of PD. Each individual factor may only contribute a slight variation and influence a small portion of the whole population. However, the potential non-linear and unpredictable interactions may exist. Overall they may account for a high proportion of the aetiological fraction.

2.6 CONCLUSIONS

This literature review supports that cigarette smoking confers a protective effect against the risk of developing PD. However, the studies reviewed indicate that none of the specific genetic polymorphisms mentioned above is unequivocally associated with PD. Since there are great discrepancies in the methodology for these studies, direct
comparison of the results can be difficult. For example, a number of studies chose idiopathic PD patients as subjects while others chose familial PD or the mixture of both; several studies attempted to select newly diagnosed cases while the majority included both incident and prevalent cases; some authors sub-divided PD cases into young- and late-onset for analyses while others did not; some studies recruited cases from hospitals or clinics while others took the effort of population-based sampling. Similar discrepancies can also be seen in control selection. The control subjects varied from healthy volunteers to spouses of cases, blood donors, or patients without neurodegenerative disorders. However, despite these methodological discrepancies, it is reasonable to believe that none of these polymorphisms may independently make a great contribution to PD in the general population, although they may be important causes among some specific population subgroups.

More importantly, some classes of chemical compounds found in cigarette smoke are substrates of these polymorphic genes. The two studies reviewed (see Section 2.5) have also suggested that there may be complex interactions between these genetic polymorphisms and these chemical compounds. However, the majority of previous genetic association studies in PD did not take into account the influence of an individual’s cigarette smoking status. It is reasonable to speculate that some gene-smoking interactions may greatly alter an individual’s susceptibility to PD, and such studies may provide more powerful analysis. Moreover, clarifying this issue and further understanding the nature of smoking’s protective effect are important since they may provide new strategies for the treatment as well as prevention of PD. To date, few known national or international studies have specifically focused on examining the potential gene-smoking interactions. Thus, this thesis has attempted to
use more powerful and reliable research approaches to examine the potential interactions between cigarette smoking and the specific genetic polymorphisms of GSTM1, GSTP1, GSTT1, GSTZ1, CYP2D6 and CYP2E1.
Table 2.1. Review of studies investigating the association between PD and various GST enzyme genes

<table>
<thead>
<tr>
<th>Study No.</th>
<th>Mutation /Allele</th>
<th>Country /Ethnicity</th>
<th>Number of Cases</th>
<th>Number of Controls</th>
<th>OR (95%CI)</th>
<th>Association with PD</th>
<th>References</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GSTM1 deletion</td>
<td>Caucasian</td>
<td>214</td>
<td>330</td>
<td>1.20 (0.85-1.70)</td>
<td>No</td>
<td>Kelada SN et al [97]</td>
<td>2003</td>
</tr>
<tr>
<td>2</td>
<td>GSTM1 deletion</td>
<td>Japanese</td>
<td>111</td>
<td>100</td>
<td>1.14 (0.75-1.72)</td>
<td>No</td>
<td>Harada S et al [174]</td>
<td>2001</td>
</tr>
<tr>
<td>3</td>
<td>GSTM1 deletion</td>
<td>Germany</td>
<td>149</td>
<td>99</td>
<td>1.02 (0.61-1.70)</td>
<td>No</td>
<td>Rahbar A et al [172]</td>
<td>2000</td>
</tr>
<tr>
<td>4</td>
<td>GSTM1 deletion</td>
<td>Sweden</td>
<td>35</td>
<td>283</td>
<td>1.23 (0.6-2.7)</td>
<td>No</td>
<td>Ahmadi A et al [73]</td>
<td>2000</td>
</tr>
<tr>
<td>5</td>
<td>GSTM1 deletion</td>
<td>Caucasian</td>
<td>167</td>
<td>225</td>
<td>n/a</td>
<td>Yes</td>
<td>Stroombergen MC et al [171]</td>
<td>1999</td>
</tr>
<tr>
<td>6</td>
<td>GSTM1 deletion</td>
<td>Australia</td>
<td>95</td>
<td>95</td>
<td>0.96 (0.54-1.69)</td>
<td>No</td>
<td>Menegon A et al [81]</td>
<td>1999</td>
</tr>
<tr>
<td>7</td>
<td>GSTM1 deletion</td>
<td>Italy</td>
<td>100</td>
<td>200</td>
<td>1.01 (0.51-2.01)</td>
<td>No</td>
<td>De Palma G et al [186]</td>
<td>1998</td>
</tr>
<tr>
<td>8</td>
<td>GSTM1 deletion</td>
<td>Caucasian</td>
<td>206</td>
<td>206</td>
<td>1.02 (0.67-1.55)</td>
<td>No</td>
<td>Nicholl et al [223]</td>
<td>1998</td>
</tr>
<tr>
<td>9</td>
<td>GSTM1 deletion</td>
<td>Caucasian</td>
<td>100</td>
<td>100</td>
<td>1.09 (0.61-1.93)</td>
<td>No</td>
<td>Bandmann O et al [173]</td>
<td>1997</td>
</tr>
<tr>
<td>Study No.</td>
<td>Mutation /Allele</td>
<td>Country /Ethnicity</td>
<td>Number of Cases</td>
<td>Number of Controls</td>
<td>OR (95%CI)</td>
<td>Association with PD</td>
<td>References</td>
<td>Year</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
<td>---------------------</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>10</td>
<td>GSTP1-104</td>
<td>Caucasian</td>
<td>214</td>
<td>330</td>
<td>1.43 (0.98-2.08)</td>
<td>No</td>
<td>No</td>
<td>Kelada SN et al [97]</td>
</tr>
<tr>
<td></td>
<td>-113</td>
<td></td>
<td></td>
<td></td>
<td>1.13 (0.73-1.75)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>GSTP1</td>
<td>Australia</td>
<td>95</td>
<td>95</td>
<td>n/a</td>
<td>Yes</td>
<td>Menegon A et al [81]</td>
<td>1999</td>
</tr>
<tr>
<td>12</td>
<td>GSTT1 deletion</td>
<td>Caucasian</td>
<td>214</td>
<td>330</td>
<td>0.85 (0.54-1.32)</td>
<td>No</td>
<td>Kelada SN et al [97]</td>
<td>2003</td>
</tr>
<tr>
<td>13</td>
<td>GSTT1 deletion</td>
<td>Australia</td>
<td>95</td>
<td>95</td>
<td>1.67 (0.82-3.39)</td>
<td>No</td>
<td>Menegon A et al [81]</td>
<td>1999</td>
</tr>
<tr>
<td>14</td>
<td>GSTT1 deletion</td>
<td>Caucasian</td>
<td>206</td>
<td>206</td>
<td>n/a</td>
<td>No</td>
<td>Nicholl et al [223]</td>
<td>1998</td>
</tr>
<tr>
<td>15</td>
<td>GSTT1 deletion</td>
<td>Italy</td>
<td>100</td>
<td>200</td>
<td>1.80 (1.03-3.15)</td>
<td>Yes</td>
<td>De Palma G et al [186]</td>
<td>1998</td>
</tr>
<tr>
<td>16</td>
<td>GSTT1 deletion</td>
<td>Caucasian</td>
<td>100</td>
<td>100</td>
<td>1.07 (0.51-2.26)</td>
<td>No</td>
<td>Bandmann O et al [173]</td>
<td>1997</td>
</tr>
<tr>
<td>17</td>
<td>GSTZ1</td>
<td>Australia</td>
<td>95</td>
<td>95</td>
<td>n/a</td>
<td>No</td>
<td>Menegon A et al [81]</td>
<td>1999</td>
</tr>
<tr>
<td>Study No.</td>
<td>Mutation /Allele</td>
<td>Country /Ethnicity</td>
<td>Number of Cases</td>
<td>Number of Controls</td>
<td>OR (95%CI)</td>
<td>Association with PD</td>
<td>References</td>
<td>Year</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
<td>---------------------</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>CYP2D6*3/*4</td>
<td>Caucasian</td>
<td>80</td>
<td>156</td>
<td>0.72 (0.45-1.19)</td>
<td>No</td>
<td>Harhangi BS et al [77]</td>
<td>2001</td>
</tr>
<tr>
<td>2</td>
<td>CYP2D6*4</td>
<td>Korean</td>
<td>93</td>
<td>121</td>
<td>6.88 (0.76-62.5)</td>
<td>No</td>
<td>Woo SI et al [89]</td>
<td>2001</td>
</tr>
<tr>
<td>3</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>139</td>
<td>113</td>
<td>1.3 (0.4-4.2)</td>
<td>No</td>
<td>Maraganore DM et al [197]</td>
<td>2000</td>
</tr>
<tr>
<td>4</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>41</td>
<td>145</td>
<td>2.1 (1.11-3.99)</td>
<td>Yes</td>
<td>Stefanovic M et al [86]</td>
<td>2000</td>
</tr>
<tr>
<td>5</td>
<td>CYP2D6*3/*4</td>
<td>Chinese</td>
<td>123</td>
<td>124</td>
<td>1.55 (0.68-3.53)</td>
<td>No</td>
<td>Ho SL et al [191]</td>
<td>1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Caucasian</td>
<td>95</td>
<td>62</td>
<td>0.90 (0.47-1.70)</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>50</td>
<td>96</td>
<td>2.2</td>
<td>Yes</td>
<td>Bon MA et al [199]</td>
<td>1999</td>
</tr>
<tr>
<td>7</td>
<td>CYP2D6PM</td>
<td>Caucasian</td>
<td>109</td>
<td>68</td>
<td>1.45 (0.60-3.51)</td>
<td>No</td>
<td>Sabbagh N et al [84]</td>
<td>1999</td>
</tr>
<tr>
<td>8</td>
<td>CYP2D6*3/*4</td>
<td>Caucasian</td>
<td>100</td>
<td>100</td>
<td>n/a</td>
<td>Yes</td>
<td>Tao E et al [224]</td>
<td>1998</td>
</tr>
<tr>
<td>9</td>
<td>CYP2D6*3/*4</td>
<td>Chinese</td>
<td>53</td>
<td>94</td>
<td>n/a</td>
<td>No</td>
<td>Lo HS et al [80]</td>
<td>1998</td>
</tr>
<tr>
<td>10</td>
<td>CYP2D6PM</td>
<td>Caucasian</td>
<td>109</td>
<td>110</td>
<td>1.20 (0.75-1.92)</td>
<td>No</td>
<td>Joost O et al [78]</td>
<td>1998</td>
</tr>
<tr>
<td>11</td>
<td>CYP2D6PM</td>
<td>Caucasian</td>
<td>100</td>
<td>200</td>
<td>14.7 (1.2-185.2)</td>
<td>Yes</td>
<td>Palma G et al [186]</td>
<td>1998</td>
</tr>
<tr>
<td>12</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>112</td>
<td>206</td>
<td>1.7 (0.94-2.45)</td>
<td>No</td>
<td>McCann SJ et al [196]</td>
<td>1996</td>
</tr>
<tr>
<td>Study No.</td>
<td>Mutation /Allele</td>
<td>Country /Ethnicity</td>
<td>Number of Cases</td>
<td>Number of Controls</td>
<td>OR (95%CI)</td>
<td>Association with PD</td>
<td>References</td>
<td>Year</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-------------------</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>13</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>47</td>
<td>47</td>
<td>2.12 (1.41-2.62)</td>
<td>Yes</td>
<td>Lucotte G et al [194]</td>
<td>1996</td>
</tr>
<tr>
<td>14</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>20</td>
<td>23</td>
<td>2.59 (0.70-9.64)</td>
<td>No</td>
<td>Kosel S et al [225]</td>
<td>1996</td>
</tr>
<tr>
<td>15</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>115</td>
<td>73</td>
<td>1.06 (0.64-1.78)</td>
<td>No</td>
<td>Gasser T et al [76]</td>
<td>1996</td>
</tr>
<tr>
<td>16</td>
<td>CYP2D6*4</td>
<td>USA</td>
<td>236</td>
<td>108</td>
<td>0.68 (0.41-1.13)</td>
<td>No</td>
<td>Sandy MS et al [85]</td>
<td>1996</td>
</tr>
<tr>
<td>17</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>272</td>
<td>720</td>
<td>n/a</td>
<td>Yes</td>
<td>Chaudhuri K et al [226]</td>
<td>1995</td>
</tr>
<tr>
<td>18</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>123</td>
<td>150</td>
<td>2.42 (1.25-4.71)</td>
<td>Yes</td>
<td>Agundez JAG et al [227]</td>
<td>1995</td>
</tr>
<tr>
<td>19</td>
<td>CYP2D6*4</td>
<td>Russia</td>
<td>38</td>
<td>70</td>
<td>1.77</td>
<td>Yes</td>
<td>Akhmedova SN et al [195]</td>
<td>1995</td>
</tr>
<tr>
<td>20</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>19</td>
<td>17</td>
<td>n/a</td>
<td>Yes</td>
<td>Rempfer R et al [228]</td>
<td>1994</td>
</tr>
<tr>
<td>21</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>105</td>
<td>105</td>
<td>0.82 (0.50-1.36)</td>
<td>No</td>
<td>Bordet R et al [229]</td>
<td>1994</td>
</tr>
<tr>
<td>22</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>50</td>
<td>110</td>
<td>1.86 (1.33-2.39)</td>
<td>Yes</td>
<td>Kurth MC et al [230]</td>
<td>1993</td>
</tr>
<tr>
<td>23</td>
<td>CYP2D6*4</td>
<td>Caucasian</td>
<td>53</td>
<td>72</td>
<td>2.70 (1.14-6.41)</td>
<td>Yes</td>
<td>Armstrong M et al [231]</td>
<td>1992</td>
</tr>
</tbody>
</table>
Table 2.3. Review of studies investigating the association between PD and CYP2E1 enzyme genes

<table>
<thead>
<tr>
<th>Study No.</th>
<th>Mutation /Allele</th>
<th>Country /Ethnicity</th>
<th>Number of Cases</th>
<th>Number of Controls</th>
<th>OR (95%CI)</th>
<th>Association with PD</th>
<th>References</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CYP2E1 RsaI</td>
<td>China</td>
<td>158</td>
<td>150</td>
<td>1.3 (0.8-2.0)</td>
<td>No</td>
<td>Wang J et al [220]</td>
<td>2000</td>
</tr>
<tr>
<td>2</td>
<td>CYP2E1 RsaI</td>
<td>UK</td>
<td>200</td>
<td>200</td>
<td>1.70 (0.59-4.88)</td>
<td>No</td>
<td>Bandmann O et al [173]</td>
<td>1997</td>
</tr>
</tbody>
</table>
CHAPTER 3: METHODOLOGY

3.1 INTRODUCTION

This thesis is comprised of three studies. A detailed description of the methodology for each specific study is provided in the relevant chapters. This chapter discusses the methodological issues that are not dealt with or only discussed in minimum detail in the manuscript chapters.

3.2 RESEARCH DESIGN

A case-only design was chosen for this thesis. The first reason for selecting this design was that it is much more efficient at detecting gene-environment interactions, the focus of this thesis. Basically, a case-only study achieves greater statistical power than a case-control study of the same sample size. Since in most situations, interactions between genetic and environmental factors such as cigarette smoking can be subtle, and recruitment of a large number of cases for a relatively rare disease like PD is usually extremely difficult, choosing a statistically powerful approach is more pragmatic.

The case-only design can be considered a transformation of the traditional case-control design. In this method, only cases are used to investigate the multiplicative effect of the exposure of interest and the susceptible genotype of interest, whereas unaffected subjects (traditionally controls) are used solely to test the independence of
the exposure and the susceptibility genotype under study. The principles of the case-only design can be derived from the traditional case-control design [103, 104].

In general, when applying a traditional case-control study to analyse gene-environment interaction, we usually display the data in a 2*4 table as follows:

Table 3.1. Gene-environmental interaction definition

<table>
<thead>
<tr>
<th>Exposure Susceptible Genotype</th>
<th>Cases (N)</th>
<th>Controls (N)</th>
<th>ORs</th>
</tr>
</thead>
<tbody>
<tr>
<td>No No</td>
<td>A1</td>
<td>B1</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>Yes No</td>
<td>A2</td>
<td>B2</td>
<td>( ORe = \frac{A2 \times B1}{A1 \times B2} )</td>
</tr>
<tr>
<td>No Yes</td>
<td>A3</td>
<td>B3</td>
<td>( ORg = \frac{A3 \times B1}{A1 \times B3} )</td>
</tr>
<tr>
<td>Yes Yes</td>
<td>A4</td>
<td>B4</td>
<td>( ORge = \frac{A4 \times B1}{A1 \times B4} )</td>
</tr>
</tbody>
</table>

Under the assumptions that ideal control subjects can be selected and both exposure and susceptibility genotype can be appropriately classified as being either ‘No’ or ‘Yes’, ORs for exposure alone (ORe), susceptible genotype alone (ORg) and the joint contribution of the genotype and exposure (ORge) can be calculated by using unexposed subjects with no exposure and no susceptible genotype as the referent group. Under an additive model, ORge can be defined as \( ORge = ORg + ORe \) while under a multiplicative model, \( ORge = ORg \times ORe \) [103, 104].
If the gene-exposure interaction exists under a multiplicative model, the OR for interaction itself (OR\textsubscript{case-control}) can be defined as OR\textsubscript{case-control} = \frac{OR_{ge}}{OR_{g}*OR_{e}}, i.e.,

\[
\]

(3.1)

From the above equation, OR\textsubscript{case-control} can also be interpreted as the OR among cases \((A1*A4)/(A2*A3))\ divided by the OR among controls \((B1*B4)/(B2*B3))\). If the test by using controls proves the independence of the exposure and susceptible genotype of interest, the OR among controls will be equal to 1 and the above equation can be simplified as:

\[
\text{OR}_{\text{case-control}} = \frac{(A1*A4)}{(A2*A3)}
\]

(3.2)

This is actually the simplified rationale for the case-only design, which means that under the required independence assumption, it is possible to obtain an estimate for interaction solely from cases [103, 104]. In this case, OR\textsubscript{case-control} can be defined as OR\textsubscript{case-only}. If OR\textsubscript{i} (either OR\textsubscript{case-control} or OR\textsubscript{case-only}) >1, it indicates more than multiplicative effects between exposure and the susceptibility genotype; if OR\textsubscript{i} <1, it indicates less than multiplicative effects. Therefore, interactive effects in this sense can be defined as significantly different frequencies of genetic factors in the exposed cases compared to the unexposed cases.
Hamajima and colleagues compared the statistical power of traditional case-control and case-only designs to detect significant gene-environment interactions by using both hypothetical (200 cases, 200 or 400 controls) and real published data (from four cancer case-control studies). The results indicated that although the two designs got similar ORs for interaction for both types of data, the case-only design got a much narrower 95%CI, i.e., the case-only design is more statistically powerful at detecting gene-environment interaction at a given study sample size [232]. Moreover, adjustment for other potential confounding variables can also be achieved by using stratified analysis, or unconditional logistic regression models, or by other means.

The second reason for selecting a case-only design was because the PD cases for the three studies were recruited over the past few years and represent a prevalence series. Hence, an appropriate comparison group for the cases for an optimal case-control study is difficult to identify and recruit. However, for a case-only design, a comparison group is not required. Unaffected subjects are used solely to examine the assumption of independence between exposure and genotypes. The collection of unaffected subjects for this purpose is much more feasible and achievable, since the differences between cases and unaffected subjects in age, gender and particularly frequency of exposure are less important. As such, the case-only approach was considered the most appropriate design to examine the gene-environment interactions in relation to the risk of developing PD for the three studies.

However, as is seen from equation 3.2, a case-only design can only calculate ORi for examining interactions. It is unable to obtain either ORg or ORe individually. Therefore, a case-only study cannot examine whether or not a genetic or
environmental factor may contribute to the risk of a disease independently. In addition, the application of a case-only design relies on the assumption that exposure is independent of genotype. Departures from this independence can have a significant impact on the results obtained using this approach [232]. This is especially crucial since it is believed that there are some genetic factors that may predispose individuals to a differential likelihood of exposure to cigarette smoke.

A case-only study can also be applied to investigate gene-environment interactions in relation to the alteration of age-at-onset of PD. In this case, the case-only design applies the same methods as the traditional case-control design, i.e., the case-only design compares the differences in age-at-onset between groups of exposed cases with and without the susceptible genotype and between groups of unexposed cases with and without the susceptible genotype. The following table provides the basis for further explanations of these differences.

Table 3.2. Gene-environment interaction in relation to age at onset of PD

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Age at onset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ever-smokers</td>
</tr>
<tr>
<td>Genotype 1</td>
<td>A</td>
</tr>
<tr>
<td>Genotype 2</td>
<td>C</td>
</tr>
</tbody>
</table>

A typical scenario might be that A is significantly greater than B, C and D while B, C and D are similar (or there may be further differences between them). We can then infer that genotype 1 interacts with cigarette smoking to confer an additional effect (positive or negative) than the effects of individual factors (i.e., genotype alone, smoking status alone, or neither) on the delay of PD onset.
To date, no study has proven a correlation between cigarette smoking and the specific genotypes under examination in this thesis. Therefore, we believed that it was important to further clarify this issue. We used a large group of unaffected subjects to examine genotype-exposure relationships. When probable relationships between particular genotypes and cigarette smoking in unaffected individuals were suspected, we precluded the case-only approach to study the interactions with those genotypes. Instead, we tentatively applied the case-control approach to analyse the data in order to generate some preliminary results about the potential gene-environment interactions. However it should be pointed out here that these results need to be interpreted cautiously with the caveat that our unaffected subjects may not represent an entirely appropriate control group for our cases.

The case-only design has been successfully used in examining gene-environment interactions in leukaemia studies [233, 234]. To date, there are no reported studies that specifically investigate gene-smoking interactions in the aetiology of PD using the case-only design.

3.3 PD CASES

PD cases were recruited from the Princess Alexandra Hospital Department of Neurology, private neurology clinics and PD support groups throughout Queensland, Australia, during the period of 1999-2004. The diagnosis of probable or definite PD was made by a movement disorders neurologist when the subject had a combination of three of the following features: resting tremor, rigidity, bradykinesia, postural
instability; or two of these features with asymmetry in tremor, rigidity or bradykinesia (a well-accepted criteria suggested by Calne and colleagues [235]). Subjects were excluded if they showed no response to L-dopa therapy, if there were features consistent with another akinetic rigid syndrome, if cognitive decline was an early feature in the presentation, or if they were unable to complete a structured questionnaire either independently or with assistance.

Standard research criteria for case eligibility for this study included: 1) white (Caucasian) ethnicity; 2) diagnosis of idiopathic PD (unknown cause); 3) physical and mental ability to complete the questionnaire; 4) provision of informed consent for participation in the study; and 5) provision of a blood sample. Finally, 400 eligible PD cases were analysed for the first two studies (Manuscripts 1 and 2) whereas 542 cases (which also included the above 400) were investigated for the third study (Manuscript 3).

The disadvantage of this method to obtain cases is that the recruited cases may not be truly representative of all cases within the specified population (i.e., Queensland, Australia). Theoretically, the most appropriate method for case selection is to recruit PD cases from the general population by random sampling, which may maximise the generalisability of the research findings. However, this was not feasible for this PhD project as PD is relatively rare in the whole population with a crude prevalence less than 0.3% in Caucasian people [1]. Also, at present, there is no population-based PD registration system from which to draw cases in Australia.
However, given that our patients were recruited from private clinics, from the public hospital system at the Princess Alexandra Hospital as well as from PD support groups throughout Queensland, they are likely to be reasonably representative and drawn from a range of socio-economic backgrounds. In addition, our primary analysis of the study cohort revealed that the age distribution and the frequencies of reported family history of PD are similar to that previously reported in population-based studies (data unpublished). Therefore, we consider that the above method for PD case recruitment is, despite its inherent drawbacks, appropriate for this research project.

**Sample Size**

The equation 3.3 suggested by Yang et al [236] was applied to calculate the required sample size for the current case-only analyses.

\[
N = \left(\frac{Z_{\alpha/2}\sqrt{V_N} + Z_{\beta}\sqrt{V_A}}{\sqrt{Z_{\alpha/2}\sqrt{V_N} + Z_{\beta}\sqrt{V_A}}}\right)\left(\log(OR_i)(\log(OR_i))\right) (3.3)
\]

Note:

- \(Z_{\alpha/2}\) and \(Z_{\beta}\): normal deviates that serve as cut-off thresholds for appropriate areas in the tails of the standard normal distribution
- \(V_N\): the variance of the logarithm of \(R_i\) under the null hypothesis
- \(V_A\): the corresponding variance under an alternative hypothesis
- \(OR_i\): the magnitude of the gene-environment interaction

Preliminary analysis of our existing data suggested that 30% to 40% of PD cases had ever smoked. We assumed that cigarette smoking conferred a protective effect with an OR of 0.5 [55, 131]. We did not assume that genetic polymorphisms themselves could
also alter the lifelong risk of developing PD. Using Equation 3.3, we calculated that a case-only study with 400 cases would have greater than 80% power to detect an interaction with ORi of 2.0 at the 5% statistical significance level (two-tailed) when the frequencies of alleles were more than 10%.

3.4 UNAFFECTED SUBJECTS

Unaffected aged subjects, who were formerly recruited as controls for other PD studies, were analysed for examining the assumption of independence between cigarette smoking and genotypes for this thesis. These unaffected subjects were recruited from three sources during the period of 1999-2004: first, 156 family-based controls consisting of spouses, siblings and care-givers of the case group; second, 246 unrelated individuals from various community groups; third, 168 randomly selected aged healthy subjects from the general population. All these subjects were recruited in Queensland, Australia, and none of them were in hospital for other related diseases at the time of data and specimen collection. The demographic characteristics of these subjects have been described previously [95] and are briefly summarised in Table 3.3.

Table 3.3. Demographic characteristics of unaffected aged subjects

<table>
<thead>
<tr>
<th>Source</th>
<th>Male (%)</th>
<th>Female (%)</th>
<th>Total (%)</th>
<th>Mean age+/-sd (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>87 (29.4)</td>
<td>69 (25.2)</td>
<td>156 (27.4)</td>
<td>69.2+/-11.3</td>
</tr>
<tr>
<td>2</td>
<td>102 (34.5)</td>
<td>144 (52.6)</td>
<td>246 (43.2)</td>
<td>65.2+/-9.7</td>
</tr>
<tr>
<td>3</td>
<td>107 (36.1)</td>
<td>61 (22.2)</td>
<td>168 (29.4)</td>
<td>66.8+/-10.4</td>
</tr>
<tr>
<td>Combined</td>
<td>296 (100)</td>
<td>274 (100)</td>
<td>570 (100)</td>
<td>67.1+/-10.6</td>
</tr>
</tbody>
</table>
402 unaffected aged subjects from the first two sources were analysed to examine the assumption of independence between cigarette smoking and genotypes for the first two studies, whereas all 570 subjects were analysed for the third study. Standard research criteria for eligibility include: 1) white (Caucasian) ethnicity; 2) physical and mental ability to complete the questionnaire; 3) provision of informed consent for participation in study; and 4) provision of a blood sample. Controls were excluded if a mental disability precluded them from completing the questionnaire.

Theoretically, for a case-only design, the most appropriate population from which to select unaffected subjects is the general population. However, unlike a case-control study, unaffected subjects in the case-only design are solely used to examine the assumption of independence between cigarette smoking and genotypes. The frequencies of different genotypes among unaffected subjects are not directly compared with those among PD cases. Therefore, this study design is much less sensitive to biases than control selection for a case-control study. Furthermore, collecting population-based unaffected subjects is time-consuming and costly. Therefore, this use of a convenient source of unaffected subjects was considered justifiable.

**Sample Size**

For this thesis, the purpose of recruiting unaffected subjects was to examine the assumption that cigarette smoking is independent of genotypes. In such a case, we could apply the principles for a case-control study to estimate the required sample size. The outcome variable was smoking status. Therefore, we considered smokers (i.e.
ever smoked) as ‘cases’ whereas non-smokers (i.e. never smoked) were ‘controls’.

The following equation was used for the calculation:

\[ N(\text{case}) = N(\text{control}) = (I^2I^*(p1^*q1+p2^*q2))/((p1-p2)(p1-p2)) \quad (3.4) \]

\[ N(\text{total}) = N(\text{case}) + N(\text{control}) \quad (3.5) \]

Note:

I: constant. At \( \alpha = 0.05 \) and \( \beta = 0.20 \) (power = 80%), \( I = 2.80 \)

p1: percentage of people with risk genotype among ‘cases’

q1: 1-p1

p2: percentage of people with risk genotype among ‘controls’

q2: 1- p2

It has been reported that all the targeted genes we examined contain proven polymorphisms with allelic frequencies greater than 5% among the general population. Therefore, the risk genotypes for all these genes were greater than 10% (for rare alleles such as those on CYP2E1 gene, we considered the heterozygotes as the risk genotypes). We therefore firstly set \( p2 \) at 10%.

Among the general population we targeted, approximately 40% of people were ever smokers and the rest (60%) had never smoked. The estimation of required sample size was adjusted for this frequency according to equation 3.6.

\[ N(\text{total}) = N(\text{case})/0.4 \quad (3.6) \]
For 80% power and at the 5% statistical significance level (two-tailed), the required sample sizes for detecting an association between cigarette smoking and genotypes are summarised in Table 3.4.

Table 3.4. Required sample sizes

<table>
<thead>
<tr>
<th>ORs</th>
<th>10%</th>
<th>20%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3</td>
<td>4423</td>
<td>1919</td>
<td>667</td>
</tr>
<tr>
<td>1.4</td>
<td>2578</td>
<td>1108</td>
<td>373</td>
</tr>
<tr>
<td>1.5</td>
<td>1706</td>
<td>726</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>1222</td>
<td>514</td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>925</td>
<td>385</td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>728</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>591</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>490</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>415</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>356</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Therefore, with 402 unaffected subjects, we had greater than 80% power to detect an association between genotypes and smoking status with ORs of less than 1.4, 1.7 and 2.2, respectively, at the 5% statistical significance level (two-tailed) when the frequencies of risk genotypes were at 40%, 20% and 10%. For the third study, the frequency of risk genotype (i.e., CYP2D6 PM) was reported to be around 8%. Based
on 570 unaffected subjects, an association with an OR around 2.0 might be detected with 80% power.

3.5 DATA COLLECTION

After informed consent was obtained from each subject, a face-to-face interview was arranged as soon as possible. Smoking and other epidemiological data were collected using a structured questionnaire (see Appendix B, C, and D) by Dr. George Mellick (one of my associate supervisors) and Ms. Coral Gartner (another PhD candidate) during the period of 1999-2004. These data were stored in an Excel database with each subject being given a unique lab identification number. Reliability tests were conducted by Ms. Coral Gartner using the Kappa statistics. These tests indicate a significant congruency of subjects’ answers to questions related to smoking habit (Kappa >0.73, p<0.00001) and other environmental exposures relevant to this thesis (all Kappa >0.56, p<0.01).

3.5.1 Smoking data

Definition of cigarette smoking: Subjects were considered smokers if they had ever smoked at least one cigarette per day for a period of a year or more before recruitment (for PD cases, prior to diagnosis).

The smoking data obtained from study subjects included:

1) the age the subject started smoking;
2) cessation history;
3) cigarettes smoked per day on average; and,
4) current smoking status.

These data allowed approximate calculation of cumulative exposure to cigarette smoke in pack-years using the following formula:

\[
\text{Pack-years} = \left( \frac{\text{cigarettes smoked per day}}{20} \right) \times \text{years smoked} \quad (3.7)
\]

**3.5.2 Other epidemiological data**

Other epidemiological data relevant to this thesis were collected, including:
1) age at which the subject was recruited for this study;
2) gender;
3) age at onset of PD in cases;
4) family history of PD: defined as first-degree relative with PD;
5) history of regular exposure to toxins: defined as weekly exposure to herbicides or pesticides for a period of six months or more.

**3.6 CANDIDATE GENES**

The first study included the genetic polymorphisms of GSTM1, GSTP1, GSTT1, and GSTZ1; the second study included CYP2E1 RsaI and PstI genetic polymorphisms; and the third study included the three common CYP2D6 genetic polymorphisms (*3, *4 and *5) which result in CYP2D6 poor metaboliser phenotypes (PMs). There are four reasons for selecting these genes. First, they are potentially involved in the
metabolism of chemical compounds in cigarette smoke; second, all these genes contain proven polymorphisms with allelic frequencies around or greater than 5%; third, all of these genetic polymorphisms have well-characterised influences on the biological activities of the respective gene products; and fourth, all of these genetic polymorphisms have been previously examined as potential genetic risk factors for PD (see Chapter 2).

3.7 DNA EXTRACTION

Peripheral blood was obtained from each PD case and unaffected subject. Blood samples are stored in a –80°C freezer. Genomic DNA was extracted by me from peripheral lymphocytes using standard methods (see Appendix E).

3.8 OPTIMISATION OF GENOTYPING ASSAYS

GSTM1, GSTP1, GSTT1, GSTZ1 AND CYP2E1 genotyping processes were performed using protocols previously published with minor modification, whereas CYP2D6 genotyping methods were mainly developed by me with assistance from associate supervisor Dr. George Mellick (details for genotyping methods are described in subsequent chapters). In each case, primer sequences, the target DNA sequence, and the length of the PCR product were checked using the human genome database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide) (provided by NCBI). Uniqueness of the target DNA sequence was confirmed by aligning the sequence to the human genome database using the BLAST function. Primer design software was also used to confirm the suitability of the primer pairs.
The reaction conditions for each specific PCR were optimised based on the following thermal cycling stages:

94°C*3min;
94°C*15sec, 61°C*30sec, 72°C*30sec, 35 cycles;
72°C*7min.

The optimisation of PCR amplification was carried out in a total volume of 25 µL PCR buffer (50 mmol/L KCl, 10 mmol/L Tris-HCl, pH9.0), containing 10ng genomic DNA, 0.2 µmol/L of each primer, 200 µmol/L of each dNTPs, 1.5 mmol/L of MgCl2 and 1.25 units of Taq polymerase. The annealing temperature was optimised by conducting the PCR reaction at gradient temperatures between 50°C and 63°C to determine the most optimal condition.

Initially, we used the DNA samples from 32 female individuals to confirm that the loci of interest are polymorphic in Caucasian Australians and to roughly estimate the polymorphic allele frequencies in this group. Examination of DNA samples from 32 females (64 alleles) has greater than 96% probability of detecting any sequence variants with allele frequency of 5% or more.

3.9 RELIABILITY AND VALIDITY OF PCR AND RFLP METHODS

Reliability and validity of the PCR and RFLP methods were assessed through re-conducting the genotype assays using at least a 10% sample of our DNA samples. In
addition, for CYP2D6 and CYP2E1, the methods were also assessed through re-conducting the genotype assays using all DNA samples with mutations (both heterozygotes and homozygotes). The results for all re-assessments were 100% concordant.

3.10 HAPLOTYPING

Haplotypes were inferred from genotype data by the implementation of the expectation maximisation algorithm using PMPLUS software [237-240]. In brief, the alleles of each bialleic marker were designated as 1 and 2, and the genotypes as 1/1, 1/2 and 2/2. For a group of individuals, their genotypes at these two loci can be summarised into a 3*3 table, with each cell in the table corresponding to a two-locus genotype. The PMPLUS software goes through each cell of the 3*3 table to obtain the haplotype counts.

3.11 DATA ANALYSIS

Definition of variables

1) Variables for evaluating the potential interaction between cigarette smoking and genetic polymorphisms are summarised in Table 3.5.
### Table 3.5. Summary of variables for interaction study

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classification</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parkinson’s disease (PD)</td>
<td>dichotomous</td>
<td>yes or no</td>
</tr>
<tr>
<td><strong>Explanatory variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td>categorical</td>
<td>never, ever or current</td>
</tr>
<tr>
<td>Pack-years of smoking</td>
<td>continuous</td>
<td></td>
</tr>
<tr>
<td>Total years of smoking</td>
<td>continuous</td>
<td></td>
</tr>
<tr>
<td>Average number of cigarettes smoked per day</td>
<td>continuous</td>
<td></td>
</tr>
<tr>
<td>GSTM1</td>
<td>dichotomous</td>
<td>deletion or no deletion</td>
</tr>
<tr>
<td>GSTP1-105</td>
<td>categorical</td>
<td>AA, AG or GG</td>
</tr>
<tr>
<td>GSTP1-114</td>
<td>categorical</td>
<td>CC, CT or TT</td>
</tr>
<tr>
<td>GSTT1</td>
<td>dichotomous</td>
<td>deletion or no deletion</td>
</tr>
<tr>
<td>GSTZ1-32</td>
<td>categorical</td>
<td>GG, AG or AA</td>
</tr>
<tr>
<td>GSTZ1-42</td>
<td>categorical</td>
<td>GG, AG or AA</td>
</tr>
<tr>
<td>CYP2E1</td>
<td>categorical</td>
<td>CC, CT or TT</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>categorical</td>
<td>EM, carrier or PM</td>
</tr>
<tr>
<td><strong>Confounding variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>continuous</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>dichotomous</td>
<td>male or female</td>
</tr>
<tr>
<td>Family history of PD (first-degree relative)</td>
<td>dichotomous</td>
<td>yes or no</td>
</tr>
<tr>
<td>Regular exposure to toxins (pesticides and/or herbicides)</td>
<td>dichotomous</td>
<td>yes or no</td>
</tr>
</tbody>
</table>

2) Variables for examining the assumption that cigarette smoking is independent of specific genotype are summarised in Table 3.5. (Classification and measurement for these variables are exactly the same as in Table 3.6.).

### Table 3.6. Summary of variables for independence of smoking and genotype
<table>
<thead>
<tr>
<th>Outcome variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking status, Pack-years of smoking, Total years of smoking, Average number of cigarette smoked per day</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Explanatory variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTM1, GSTP1-105, GSTP1-114, GSTT1, GSTZ1-32, GSTZ1-42, CYP2E1, CYP2D6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Confounding variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, gender, Family history of PD, Regular exposure to toxins</td>
</tr>
</tbody>
</table>

**Bivariate analyses**

For the comparison of continuous variables, we firstly checked the assumption that they were normally distributed. If the distribution was normal, results were expressed as means and standard deviations (sd) and a classical t-test or one-way ANOVA was used. If the distribution was skewed, then a logarithmic transformation was attempted to normalise the data. If successful, geometric means were used to summarise the data and again classical t-test or one-way ANOVA was used. If not transformable, results were expressed as medians and range (minima, maxima), and a Mann-Whitney test was used.

For the comparison of dichotomous or categorical variables, they were summarised as a table with counts and percentages, and the associations were tested using Pearson’s chi-squared test.

**Multivariate analyses:**
This case-only approach represents one type of observational study for which the levels of factors could not be controlled by the researchers. In such a case, a strong potential for confounding effects does exist. Therefore, crude bivariate analyses were adjusted for potential confounding effects of the variables listed earlier. Unconditional logistic regression was used to model the dichotomous outcome variable while a multivariate linear regression model was used to model the outcome of continuous variables. Then two types of inferential tests using logistic regression (tests of models and tests of individual predictors) were applied. The unadjusted and adjusted estimates were compared, and if different by more than 10% between any of the estimates, the variables listed earlier were considered to be potential confounding factors, and the adjusted results were reported.

1) To evaluate the potential interaction between cigarette smoking and genetic polymorphisms in relation to lifelong risk of PD with a case-only design, logistic regression directly generated OR_{case-only} that provides a measure of the multiplicative interactive effect.

2) To examine the assumption that cigarette smoking is independent of specific genotype, logistic regression generated an OR that indicates the magnitude of the relationship and whether it was significantly different from the null of 1.0.

3) To investigate the potential interaction between cigarette smoking and genetic polymorphisms in relation to the alteration of age-at-onset of PD, multiple linear regression was used to calculate the average age-at-onset for different combinations of genotype and smoking status.
All results involving genotypes were adjusted for age (gene-smoking independence test) or age-at-onset (gene-smoking interaction test), gender, and family history of PD. Results involving haplotype data were not adjusted because haplotypes were inferred from genotype data by the implementation of the expectation maximisation algorithm [237-240].

**Hardy-Weinberg Equilibrium**

Deviation in the genotype counts from Hardy-Weinberg equilibrium was tested using the Chi-square test for goodness of fit.

**Assumptions**

For the bivariate t-tests or one-way ANOVA of continuous variables, the explanatory variable should be approximately normal. The distribution was considered approximately normal if all the following conditions were met:

- mean was within 10% of median
- the sd was less than or equal to half the mean
- the mean ± 3 sd approximated the minimum and maximum
- the skewness coefficient was within ± 3
- the kurtosis coefficient was within ± 3
- the histogram looked approximately symmetrical and bell-shaped
The similarity of variances (standard deviations) across comparison groups was examined using Levene’s test. When the variances were similar, Duncan’s test was applied; when the variances were significantly different, Dunnett’s test was applied.

For Pearson’s chi-square tests that were used for bivariate analyses of categorical (including dichotomous) variables, it was considered valid when expected cell sizes were 5 or more. This was ensured by collapsing categories where required and meaningful, or excluding categories where not meaningful.

Theoretically, the logistic regression used for multivariable analyses is relatively free of limitations. The distributions of the independent variables do not need to be normally distributed, linearly related, or of equal variance within each group. But if too many cells are with no cases, it is not acceptable. However, this did not occur in our case.
Significance level to be applied to analyses

All hypotheses were tested with two-tailed tests, and statistical significance was declared at the conventional p<0.05 level. All estimates were quoted with corresponding 95% confidence intervals (95%CI).

3.12 ETHICS STATEMENT

Ethical Approvals associated with this thesis work were granted by QUT University Human Research Ethics Committee (QUT Ref No 2510H), by UQ University Human Research Ethics Committee, and by Princess Alexandra Hospital Human Research Ethics Committee. The ethical approvals covered:

- Recruitment of PD cases and unaffected, healthy subjects;
- Collection of data with the questionnaires;
- Access to medical records about patients participating in the study (with the patient’s written consent) to confirm date of diagnosis and presenting symptom;
- Collection of a 10mL sample of blood from the participants for the purpose of genetic analysis; and
- Examination of genetic polymorphisms using DNA extracted from the blood samples collected.
3.13 HEALTH AND SAFETY APPROVAL

Health and Safety Approval for the project was also granted by the QUT School of Public Health Workplace Health and Safety Officer, and by the QUT Faculty of Health Workplace Health and Safety Officer.

*Correspondence confirming the above-mentioned approvals is located in Appendix A.*
CHAPTER 4: CASE-ONLY STUDY OF INTERACTIONS BETWEEN GENETIC POLYMORPHISMS OF GSTM1, P1, T1 AND Z1 AND SMOKING IN PARKINSON’S DISEASE

Citation:

Deng YF, Newman B, Dunne MP, Silburn PA and Mellick GD: Case-only study of interactions between genetic polymorphisms of GSTM1, P1, T1 and Z1 and smoking in Parkinson’s disease. Neuroscience Letters. 366(3):326-331, 2004

Contribution of authors:

The candidate performed all genotyping, data analysis and wrote the entire manuscript. Other authors contributed to the manuscript in terms of providing feedback on the analyses and initial drafts.

Note to examiners:

Tables containing results from additional analyses mentioned in the published paper, but for which data were not shown, are included at the end of this Chapter.
ABSTRACT

Current opinion contends that complex interactions between genetic and environmental factors play a role in the etiology of Parkinson’s disease (PD). Cigarette smoking is thought to reduce risk of PD, and emerging evidence suggests that genetic factors may modulate smoking’s effect. We used a case-only design, an approach not previously used to study gene-environment interactions in PD, specifically to study interactions between glutathione S-transferase (GST) gene polymorphisms and smoking in relation to PD. 400 PD cases (age at onset=60.0+/−10.7 years) were genotyped for common polymorphisms in GSTM1, P1, T1 and Z1 using well-established methods. Smoking exposure data were collected in face-to-face interviews. The independence of the studied GST genotypes and smoking exposure was confirmed by studying 402 healthy, aged individuals. No differences were observed in the distributions of GSTM1, T1 or Z1 polymorphisms between ever-smoked and never-smoked PD cases using logistic regression (all p>0.43). However, GSTP1 *C haplotypes were over-represented among PD cases who ever smoked (odds ratio for interaction (ORi) = 2.00 (95% CI: 1.11-3.60, p=0.03)). Analysis revealed that ORi between smoking and the GSTP1-114Val carrier status increased with increasing smoking dose (p=0.02 for trend). These data suggest that one or more GSTP1 polymorphisms may interact with cigarette smoking to influence the risk for PD.

Keywords: glutathione S-transferase, genetic polymorphism, interaction, Parkinson’s disease, case-only study
4.1 INTRODUCTION

Current opinion contends that the etiology of idiopathic Parkinson’s disease (PD) is likely to be multifactorial, and that complex interactions between genetic susceptibility and environmental exposures play a role [241]. Cigarette smoking is less common among PD cases [242]. However, the underlying mechanism for this apparent protective effect is unclear. Evidence suggests that genetic factors may modulate this effect [92, 94, 243, 244], but few studies have focused explicitly on assessing these gene-smoking interactions in PD.

It is important to clarify this issue to help uncover environmental and genetic factors related to PD. Moreover, understanding such interactions may help define neuroprotective substrates or pathways that could be exploited as novel prevention or treatment strategies for PD. If environmental factors such as chemical compounds in cigarette smoke influence PD risk, differences in the ability of individuals to metabolize these substances may alter their susceptibility to the disease.

Glutathione-S-transferases (GSTs) consist of a super-family of dimeric phase 2 metabolic enzymes that catalyse the conjugation of reduced glutathione with electrophilic groups of a wide variety of compounds. They play an important role in detoxifying cytotoxic agents and protecting cellular macromolecules. In some cases, glutathione conjugation can also result in chemical intermediates that are more reactive than the parent compounds. In humans, GSTs are divided into major classes that have distinct substrate specificities and tissue distributions. Several GST genes (M1, P1, T1, and Z1) exhibit polymorphic traits that directly alter gene function due
to the altered quantity or activity of enzymes caused by nucleotide variation or deletion in the genes [245].

A number of case-control studies have investigated associations between GST genetic polymorphisms and PD risk, but few have considered the influence of smoking on this genetic association. Palma and colleagues [246] applied stratified analyses to examine the influence of GSTM1 deletion on associations between smoking and PD risk. They noticed that the protective effect of cigarette smoking was lost among patients with GSTM1-deleted genotype. A recent study by Kelada and colleagues [97] also used stratified analyses to examine the influence of smoking on associations between GSTP1 genetic polymorphisms and PD risk. They observed an increased frequency of GSTP1-Ile105Val heterozygotes among PD cases, compared to controls, in subjects who ever smoked (OR=1.92) [97]. Taken together, these findings suggest that a focused examination of GST gene-smoking interactions in PD is warranted.

The current study specifically aimed to examine whether functional genetic variations in GST genes (M1, P1, T1, and Z1) interact with smoking to influence the development of PD. We applied a case-only design based on the following considerations: First, PD is a relatively rare disease and a case-only design is considered the most powerful method for studying gene-environment interactions; it achieves greater statistical power than a case-control study of the same sample size [103, 104]. Second, for an optimal case-control study using our cases, an appropriately matched control group is difficult to ascertain. However, for a case-only study, affected subjects are used to investigate the multiplicative effects of the exposures and susceptible genotypes of interest. Non-case subjects (traditionally
controls) are used to test the independence between the exposure and the susceptible genotype. When non-case subjects are used for this purpose, strict matching to cases is not required and many of the limitations inherent in the case-control approach are minimised [103]. We know of no previous reports of studies that have specifically used this method to investigate gene-smoking interactions in the etiology of PD.

4.2 MATERIALS AND METHODS

Subjects

400 eligible PD cases were recruited from the Princess Alexandra Hospital Department of Neurology and private neurology clinics in Queensland, Australia. The diagnosis of probable or definite PD was made by a movement disorders neurologist when the subject had a combination of three of the following features: resting tremor, rigidity, bradykinesia, postural instability; or two of these features with asymmetry in tremor, rigidity or bradykinesia. Subjects were excluded if they showed no response to L-dopa therapy, if there were features consistent with another akinetic rigid syndrome, if cognitive decline was an early feature in the presentation, or if they were unable to complete a structured questionnaire either independently or with assistance. 402 unaffected, aged subjects who were formerly recruited as controls for other PD studies [247-251] were analysed to examine the assumption of independence between cigarette smoking and genotypes. These subjects were white Caucasian and consisted of 156 spouses of patients and 246 other volunteers recruited from patient neighborhoods and from community organizations. They were able to complete a structured questionnaire either independently or with assistance. All participants gave appropriate informed consent and provided blood specimens for the purposes of DNA
extraction. The project was approved by human research ethics committees at each of the participating institutions.

Smoking and other epidemiological data were collected in face-to-face interviews using a structured questionnaire. Subjects were considered smokers if they had ever smoked at least one cigarette per day for a period of a year or more. The smoking data obtained from study subjects included: the age the subject started smoking; cessation history; cigarettes smoked per day during this time period; and current smoking status. These data allow calculation of cumulative exposure to cigarette smoke in pack-years. Our questionnaire also collected data pertaining to family history of PD (first-degree relatives with PD), age (or age at onset of PD in cases), history of regular exposure to toxins (herbicides or pesticides, defined by weekly exposure for a period of six months or more).

Genotyping
Genomic DNA was extracted from peripheral lymphocytes using standard methods. GSTM1 and GSTT1 whole gene deletions were detected simultaneously in a single assay using a multiplex PCR approach adapted from [252] with the beta-globin gene being co-amplified as an internal control. Four amino acid altering single nucleotide polymorphisms (SNPs) in the GSTP1 and GSTZ1 genes were examined using previously published methods [81] with minor changes. These SNPs are listed in Table 2 together with their reference SNP identification numbers (refSNP ID) from the NCBI SNP database. Primer sequences, annealing temperatures, digest conditions and fragment lengths are available from the authors on request. Reliability and validity of the PCR and RFLP methods were assessed through re-conducting the
genotype assays using a 10% random sample of DNA samples. The results were 100% concordant. GSTP1 and GSTZ1 haplotypes were inferred from genotype data by the implementation of the expectation maximization algorithm using PMPLUS software.

**Statistical analysis**

Differences in genotype and haplotype frequencies were assessed using Pearson’s chi-squared test. Case-only analyses compared genotype frequencies in PD cases between smokers and non-smokers to obtain ORi (ORi\text{case-only}). For comparison purposes we also analysed our data using a traditional case-control approach to obtain ORi (ORi\text{case-control}), which can be defined as $\text{ORi}_{\text{case-control}} = \frac{\text{ORge}}{\text{ORg \ast ORe}}$. ORg is the odds ratio comparing genotype between cases and controls who do not expose to the environmental risk factor; ORe is the odds ratio comparing environmental factor between cases and controls who do not carry the genetic risk factor; and ORge is the odds ratio comparing cases who are jointly exposed to both genetic and environmental risk factors with controls who are exposed to neither [104]. All odds ratios involving genotypes were calculated by logistic regression, adjusted for age at onset, gender, family history of PD and toxin exposure. Odds ratios involving haplotype were not adjusted because haplotypes were inferred from genotype data. ORi>1 indicates more than multiplicative effects between gene and exposure; ORi=1 implies solely multiplicative effects; and ORi<1 suggests less than multiplicative effects.
4.3 RESULTS

The characteristics of our PD cases are shown in Table 1. There were 57.5% males and 42.5% females with the average onset age at 60.0 years (sd=10.7 years). The observed smoking rate (ever-smokers) was 43.2%, which is much lower than the expected smoking rate of 56.3% (based on Australian National Health Survey data, 1995, adjusted to the older distribution of our cases). This yields an OR for smoking and PD of 0.6, suggesting that the lifetime smoking status among our PD cases is lower than the general population, a finding consistent with other studies [242].

In both PD cases and unaffected, aged people, the genotype distributions for all loci were in Hardy-Weinberg equilibrium (p>0.39 and p>0.30, respectively). Among our unaffected, aged subjects (mean age: 63.9 years, sd: 11.4 years), the genotype frequencies at each locus were similar to those reported in other Caucasian populations. There were no relationships between genotypes and the age at which individuals started smoking or the number of pack-years smoked (p>0.27). Therefore, these polymorphisms appear to be independent of smoking in unaffected, aged individuals.

Case-only analyses of the different genotypes at the various loci among the PD group of ever- and never-smokers are summarized in Table 2. Overall, we found no differences in genotype frequencies between the ever- and never-smoked groups (all p-values >0.15). However, analysis of the GSTP1-Ala114Val locus revealed that the proportion of carriers of the 114Val allele (mutant) increased with increasing smoking dose from 0 to >30 pack-years. Homozygotes of the 114Ala allele (wild-type)
decreased with increasing smoking dose (trend test: p=0.02, Table 3). This trend existed both in male and female cases (data not shown). This dose-effect relationship was most significant in the group of cases with late-onset PD (i.e., age at onset > 55 years) with the ORi\textsubscript{case-only} values of 1.88 (95%CI: 0.65-5.48) and 2.63 (95%CI: 1.07-6.49) for >0-10 and >10 pack-years, respectively. No similar trend was found among our unaffected, aged subjects (p=0.42).

Haplotype analyses revealed significant differences for GSTP1 haplotypes between smoking and non-smoking PD cases (ORi\textsubscript{case-only} for *C haplotype=2.00 (95%CI: 1.11-3.60), p=0.03). In this case, smoking-exposed PD cases were more likely to posses the *C haplotype defined by A to G base-pair transition at nucleotide +313 and C to T base-pair transition at nucleotide +341 (at amino acid level, valine at both positions 105 and 114).

4.4 DISCUSSION

We conducted a case-only study to analyse specifically whether interactions between polymorphisms of GST genes (M1, P1, T1 and Z1) and smoking influence the development of PD. Our analysis points to an interaction between the GSTP1 gene and smoking that may influence PD risk. Wild-type GSTP1 (114Ala homozygote) tended to be less common whereas 114Val carriers (mutant) tended to be more common in smoking-exposed PD cases with an apparent dose-dependent trend (p=0.02). Haplotype analysis also indicates that there is a significantly different distribution for the *C haplotype (base-pair transitions at both nucleotides +313 and +341) between smoking and non-smoking PD cases (9.1% vs 4.4%, p=0.03). Since
there was no similar trend among unaffected, aged subjects (p = 0.42), this finding suggests there is a GSTP1 gene-smoking interaction that influences the development of PD. This trend was more pronounced in PD cases of later onset (>55 years). The age-dependent effect may reflect a more prominent environmental component to late-onset PD compared to early-onset cases. Recent twin studies also suggest that early-onset PD is more likely to involve rare genetic mutations with a more direct influence on disease status [253].

Our study specifically focused on using a case-only design to investigate gene-smoking interactions. While this approach has the advantage of eliminating the biases introduced by using a conventional control group, a disadvantage is that the case-only design is unable to examine the potential associations between each individual factor (either environmental or genetic) and PD risk [104]. This constrains our examination of interactive effects to a multiplicative model and, in this case, precludes definite information as to whether polymorphisms in these GST genes (M1, P1, T1 and Z1) play individual roles in PD in the absence of cigarette smoking. However, we are able to comment on this if our data are analysed using a conventional case-control design. This must be done cautiously with the caveat that our unaffected subjects may not represent an entirely appropriate control group for our cases. Such an analysis revealed no significant associations between GSTM1 or GSTZ1 genotypes and PD risk. However, GSTT1 null individuals were slightly over-represented in the PD group compared to controls: OR=1.54 (95%CI: 1.05-2.24, p=0.03). With respect to GSTP1, there was no difference in 105Val carriers between cases and controls (OR=0.89, 95%CI: 0.65-1.23). These ORs were not significantly influenced by stratifying the sample for smoking status.
GSTP1-114Val carriers were less common among PD cases compared to controls (OR=0.58, 95%CI: 0.38-0.90, p=0.02). Haplotype analysis also showed that the frequency of *C haplotypes was marginally lower in PD cases compared to controls (6.4% vs 8.9%), although this did not reach statistical significance (OR= 0.71, 95% CI: 0.47-1.05, p=0.09). Calculations of OR_{case-control} between smoking and GSTP1-114Val carriers using a case-control approach provide some support for the idea that true interactive effects exist. When our data were stratified for the 114 genotype, we found the OR for smoking was 2.01 (0.91-4.45) for individuals who carry the 114Val allele. The corresponding OR_{case-control} values were of similar magnitude and in the same direction to those obtained using the case-only design (footnotes in Tables 2 and 3). However, these values did not reach statistical significance. This probably reflects the increased power of the case-only design over similarly sized case-control studies [103, 104]. Nonetheless, if this is a real result, the effect size of these interactions is relatively small. Given that we performed no correction of significance values for multiple comparisons, it is also possible that the result is a chance finding.

GSTP1 is expressed in human brain and the blood-brain barrier. Given that the polymorphisms we examined have functional significance, it is possible that altered GSTP1 gene function may influence the development of neurodegenerative diseases including PD. Our case-only results can be described in two alternative ways (although these need not be mutually exclusive): 1) wild-type GSTP1-114 homozygote smokers are less likely to develop PD; or 2) GSTP1 variant carriers (114Val carriers) who smoke are at increased risk for PD. Analysis of our data by conventional case-control design supports the latter alternative. Future study is
required to clarify this issue. This is of particular interest because smoking is commonly considered to be protective against the development of PD.

One possible explanation for our data is that the protective effect of particular compounds in cigarette smoke requires activation by normal function of the GSTP1 gene. Loss or significant decrease of GSTP1 gene function would lead to a loss of the protective effect. Alternatively, in the absence of normal GSTP1 function, smoking could, in fact, increase risk of PD. Cigarette smoking is known to generate oxidative agents both directly, through metabolism of the components of cigarette smoke, and indirectly, through nicotine-induced activation of dopaminergic activity. GSTP1-114Val carriers may lack the ability to detoxify this increased load of reactive oxygen species. Similar to our results, a recent case-control study by Kelada and colleagues suggested that genetic polymorphism in GSTP1 might interact with smoking and influence the development of PD [97]. Their stratified analysis revealed a significant association for heterozygotes at the Ile105Val locus in smokers with no difference seen between never-smoked cases and controls. Analysis of our data in an identical way did not replicate their result. However, it is of interest that if Kelada’s data are analysed using a case-only approach, the OR_{case-only} obtained for GSTP1-105Val carriers is 1.62 (0.91-2.90), which is similar in magnitude and direction to our GSTP1-114 locus finding. Unfortunately, the corresponding data for the Ala114Val locus were not provided [97]. There is a current paucity of data on the differential effects of the corresponding haplotypes on GSTP1 function, particularly for the wide variety of potential GST substrates that are constituents or metabolites of cigarette smoke. Further work is required in this area to clarify how GSTP1 gene-smoking interactions influence the risk for PD.
A limitation of the case-only approach is the requirement that the examined genotypes and exposures are independent of one another in unaffected individuals. Departures from this independence can have a significant impact on the results obtained using this approach [254]. However, in the current study such effects are unlikely to influence our conclusions, given that we demonstrated independence between GSTP1-Ala114Val genotype and smoking in a relatively large number of unaffected, aged subjects (n=402) (OR=1.10, 95%CI: 0.61-2.00, p=0.74). This assured the appropriate application of a case-only design in our study. Moreover, given the number of PD cases (n=400), the case-only approach provided sufficient statistical power to detect modest interactive effects between genotype and smoking status.

To our knowledge, this is the first genetic epidemiological study of PD using a case-only design to address specifically the issue of gene-smoking interactions. Our data suggest: 1. smoking exposure is independent of GSTM1, P1, T1 and Z1 genotypes; 2. there are no multiplicative interactive effects linking smoking and GSTM1, T1 or Z1 genotypes with the risk for PD; 3. there is a multiplicative interactive effect between smoking and GSTP1 haplotype – particularly for genotypes carrying the 114Val allele.

ACKNOWLEDGEMENTS

This work was supported in part by Parkinson’s Queensland Inc., the Brain Foundation of Australia and the Geriatric Medical Foundation of Queensland. The authors would like to thank Dr. Diana Battistutta, Coral Gartner and Zaimin Wang for
assistance with data analysis, and Daniel D. Buchanan for assistance with PCR optimization.
Table 4.1. Characteristics of PD cases

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>230</td>
<td>57.5%</td>
</tr>
<tr>
<td>Female</td>
<td>170</td>
<td>42.5%</td>
</tr>
<tr>
<td><strong>Age at onset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;=50</td>
<td>79</td>
<td>19.8%</td>
</tr>
<tr>
<td>51-60</td>
<td>121</td>
<td>30.4%</td>
</tr>
<tr>
<td>61-70</td>
<td>130</td>
<td>32.7%</td>
</tr>
<tr>
<td>71-80</td>
<td>63</td>
<td>15.8%</td>
</tr>
<tr>
<td>&gt;80</td>
<td>5</td>
<td>1.3%</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never-smoked</td>
<td>227</td>
<td>56.8%</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>159</td>
<td>39.8%</td>
</tr>
<tr>
<td>Current smoker</td>
<td>14</td>
<td>3.4%</td>
</tr>
<tr>
<td><strong>Pack-years smoked+</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>227</td>
<td>57.6%</td>
</tr>
<tr>
<td>&gt;0-10</td>
<td>69</td>
<td>17.5%</td>
</tr>
<tr>
<td>&gt;10-20</td>
<td>28</td>
<td>7.1%</td>
</tr>
<tr>
<td>&gt;20-30</td>
<td>26</td>
<td>6.6%</td>
</tr>
<tr>
<td>&gt;30</td>
<td>44</td>
<td>11.2%</td>
</tr>
</tbody>
</table>

+Six subjects with smoking status did not provide detailed information on smoking.
Table 4.2. Genotype frequencies among never- and ever-smokers in PD cases

<table>
<thead>
<tr>
<th></th>
<th>Ever-</th>
<th>Never-</th>
<th>OR (_\text{case-only}) (95%CI)+</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GSTM1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1</td>
<td>89 (51.4%)</td>
<td>103 (46.2%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>*0</td>
<td>84 (48.6%)</td>
<td>120 (53.8%)</td>
<td>0.87 (0.57-1.33)</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>GSTT1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1</td>
<td>135 (80.0%)</td>
<td>176 (80.0%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>*0</td>
<td>34 (20.0%)</td>
<td>44 (20.0%)</td>
<td>0.99 (0.58-1.69)</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>GSTP1-105 (rs947894)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>74 (43.0%)</td>
<td>96 (42.3%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>79 (45.9%)</td>
<td>102 (44.9%)</td>
<td>0.96 (0.61-1.50)</td>
<td>0.85</td>
</tr>
<tr>
<td>GG</td>
<td>19 (11.1%)</td>
<td>29 (12.8%)</td>
<td>0.78 (0.39-1.55)</td>
<td>0.48</td>
</tr>
<tr>
<td>AG+GG</td>
<td>98 (57.0%)</td>
<td>131 (57.7%)</td>
<td>0.92 (0.60-1.41)</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>GSTP1-114 (rs1799811)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>143 (83.2%)</td>
<td>203 (89.8%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>26 (15.1%)</td>
<td>23 (10.2%)</td>
<td>1.42 (0.75-2.67)</td>
<td>0.28</td>
</tr>
<tr>
<td>TT</td>
<td>3 (1.7%)</td>
<td>0 (0%)</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>CT+TT</td>
<td>29 (16.8%)</td>
<td>23 (10.2%)</td>
<td>1.57 (0.85-2.92)*</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>GSTZ1-32 (rs3177427)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>77 (44.8%)</td>
<td>103 (46.6%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>76 (44.2%)</td>
<td>93 (42.1%)</td>
<td>1.02 (0.65-1.60)</td>
<td>0.92</td>
</tr>
<tr>
<td>AA</td>
<td>19 (11.0%)</td>
<td>25 (11.3%)</td>
<td>1.25 (0.61-2.53)</td>
<td>0.54</td>
</tr>
<tr>
<td>AG+AA</td>
<td>95 (55.2%)</td>
<td>118 (53.4%)</td>
<td>1.06 (0.70-1.63)</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>GSTZ1-42 (rs3177429)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>139 (83.2%)</td>
<td>181 (83.0%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>26 (15.6%)</td>
<td>36 (16.5%)</td>
<td>0.79 (0.44-1.42)</td>
<td>0.43</td>
</tr>
<tr>
<td>AA</td>
<td>2 (1.2%)</td>
<td>1 (0.5%)</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>AG+AA</td>
<td>28 (16.8%)</td>
<td>37 (17.0%)</td>
<td>0.85 (0.48-1.51)</td>
<td>0.58</td>
</tr>
</tbody>
</table>

\(^+\text{ORi: odds ratios for interactive effects between genotype and smoking status, adjusted for gender, age at onset, family history of PD and toxin exposure.}\)

\(^*\text{ORi case-control}=1.60 (0.68-3.79), \ p=0.28 \text{ (using a conventional case-control approach).}\)
Table 4.3. Pack-years and GSTP1-114 genotypes among PD cases

<table>
<thead>
<tr>
<th>Pack-years</th>
<th>CC</th>
<th>CT+TT</th>
<th>ORi\textsubscript{case-only} (95%CI)+*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>203 (89.8%)</td>
<td>23 (10.2%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>&gt;0-10</td>
<td>61 (88.4%)</td>
<td>8 (11.6%)</td>
<td>1.04 (0.42-2.59)</td>
<td>0.93</td>
</tr>
<tr>
<td>&gt;10-30</td>
<td>44 (81.5%)</td>
<td>10 (18.5%)</td>
<td>1.74 (0.74-4.04)</td>
<td>0.20</td>
</tr>
<tr>
<td>&gt;30</td>
<td>34 (79.1%)</td>
<td>9 (20.9%)</td>
<td>2.13 (0.86-5.23)</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Trend for increasing ORi with increasing dose 0.02

+ORi: odds ratios for interactive effects between genotype and smoking status, adjusted for gender, age at onset, family history of PD and toxin exposure. *ORi\textsubscript{case-control} were 1.40 (0.36-5.48), 1.13 (0.34-3.77), 2.27 (0.64-8.08) for >0-10, >10-30 and >30 pack-years, respectively.
Table A-4.1 lists SNPs together with their reference SNP identification numbers (refSNP ID) (http://www.ncbi.nlm.nih.gov/). This Table also presents the primer sequences and endonucleases we used for PCR and RFLP reactions. +For GSTM1, GSTT1 and Beta-globin, listed are frequencies of deletion (*0) and non-deletion (*1) of the loci.
Table A-4.2 Independence of smoking status and genotypes in unaffected individuals

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genotypes</th>
<th>ORs (95%CI)+</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTM1 *1</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTM1 *0</td>
<td>1.23 (0.74-2.05)</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>GSTT1 *1</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTT1 *0</td>
<td>0.90 (0.52-1.57)</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>GSTP1-105A</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTP1-105A</td>
<td>1.46 (0.90-2.38)</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>GSTP1-105A</td>
<td>1.30 (0.64-2.66)</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>GSTP1-105A</td>
<td>1.43 (0.90-2.26)</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>GSTP1-114C</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTP1-114C</td>
<td>1.18 (0.64-2.17)</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>GSTP1-114C</td>
<td>0.41 (0.04-4.15)</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>GSTP1-114C</td>
<td>1.10 (0.61-2.00)</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>GSTZ1-32G</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTZ1-32G</td>
<td>0.87 (0.53-1.43)</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>GSTZ1-32G</td>
<td>0.66 (0.29-1.49)</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>GSTZ1-32G</td>
<td>0.83 (0.52-1.33)</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>GSTZ1-42G</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTZ1-42G</td>
<td>0.97 (0.52-1.82)</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>GSTZ1-42G</td>
<td>0.29 (0.03-2.89)</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>GSTZ1-42G</td>
<td>0.89 (0.49-1.64)</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>

+ ORs: odds ratios, adjusted for age, gender, and family history of PD.

Table A-4.2 presents data demonstrating the independence of genotypes and smoking status in unaffected individuals. Among unaffected, aged subjects (mean age: 63.9 years, sd: 11.4 years), the genotype frequencies at each locus were similar to those reported in other Caucasian populations. There were no significant associations between genotypes and smoking status. Caution must be taken for the GSTP1-105 locus since the p-value is 0.12. However, considering our sample size and its relatively high frequency, we made the decision of no association.
Table A-4.3 Age at onset, pack-years and GSTP1-114 genotypes among PD cases

<table>
<thead>
<tr>
<th>Age</th>
<th>Pack-years</th>
<th>CC</th>
<th>CT+TT</th>
<th>ORi (95%CI)+</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=55</td>
<td>0</td>
<td>58 (58.0%)</td>
<td>11 (64.7%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>&gt;0-10</td>
<td>22 (22.0%)</td>
<td>1 (5.9%)</td>
<td>n/a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>20 (20.0%)</td>
<td>5 (29.4%)</td>
<td>1.17 (0.34-4.00)</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>&gt;55</td>
<td>0</td>
<td>143 (59.6%)</td>
<td>12 (36.4%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>&gt;0-10</td>
<td>39 (16.3%)</td>
<td>7 (21.2%)</td>
<td>1.88 (0.65-5.48)</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>58 (24.2%)</td>
<td>14 (42.4%)</td>
<td>2.63 (1.07-6.49)</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>

+ORi: odds ratios for interactive effects between gene and smoking, adjusted by gender, age at onset, familial history and pesticide exposure.

Table A-4.3 presents age-at-onset, pack-years and GSTP1-114 genotypes among PD cases. As stated in the manuscript, the proportion of carriers of the 114Val allele (mutant genotype) increased with increasing smoking dose from 0 to >30 pack-years whereas homozygotes of the 114Ala allele (wild genotype) decreased with increasing smoking dose (Table 4.2.). Table A-4.3 shows that this dose-effect relationship was most significant in the group of cases with late-onset PD (i.e., AAO > 55 years), with the ORi case-only values of 1.88 (95%CI: 0.65-5.48) and 2.63 (95%CI: 1.07-6.49) for >0-10 and >10 pack-years, respectively. Table A-4.3 also shows that no similar trend was found among our unaffected, aged subjects (p=0.42).
### Table A-4.4 Haplotype frequencies among never- and ever-smokers in PD cases

<table>
<thead>
<tr>
<th>Haplotypes+</th>
<th>Ever-smokers</th>
<th>Never-smokers</th>
<th>ORi (95%CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GSTP1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*A</td>
<td>225 (66.0%)</td>
<td>290 (64.6%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>*B</td>
<td>85 (24.9%)</td>
<td>139 (31.0%)</td>
<td>0.79 (0.57-1.09)</td>
<td>0.17</td>
</tr>
<tr>
<td>*C</td>
<td>31 (9.1%)</td>
<td>20 (4.4%)</td>
<td>2.00 (1.11-3.60)</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>GSTZ1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*A</td>
<td>225 (67.3%)</td>
<td>293 (67.7%)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>*B</td>
<td>79 (23.7%)</td>
<td>103 (23.8%)</td>
<td>1.00 (0.71-1.40)</td>
<td>0.94</td>
</tr>
<tr>
<td>*C</td>
<td>30 (9.0%)</td>
<td>37 (8.5%)</td>
<td>1.06 (0.63-1.76)</td>
<td>0.96</td>
</tr>
</tbody>
</table>

+Haplotypes: for GSTP1, *A haplotype defined by no base-pair transition at nucleotides +313 and +341 (at amino acid level, Ile at position 105 and Ala at position 114); *B haplotype defined by an A to G base-pair transition at nucleotide +313 (at amino acid level, Val at position 105); *C haplotype defined by A to G base-pair transition at nucleotide +313 and C to T base-pair transition at nucleotide +341 (at amino acid level, Val at both positions 105 and 114); for GSTZ1, *A haplotype defined by no base-pair transition (at amino acid level, Lys at position 32 and Arg at position 42); *B haplotype defined by an A to G transition (at amino acid level, Glu at position 32); *C haplotype defined by A to G base-pair transitions at both nucleotides (at amino acid level, Glu at position 32 and Gly at position 42).

Table A-4.4 presents the haplotype frequencies for GSTP1 and GSTZ1 among never- and ever-smokers in PD cases. As expected, the algorithm for inferring haplotypes generated alleles within the same gene in strong linkage disequilibrium with each other. Significant differences were observed for GSTP1 haplotypes between smoking and non-smoking PD cases (ORi\textsubscript{case-only} for *C haplotype = 2.00 (95%CI: 1.11-3.60), p = 0.03). In this case, smoking-exposed PD cases were more likely to possess the *C haplotype defined by A to G base-pair transition at nucleotide +313 and C to T base-pair transition at nucleotide +341 (at amino acid level, valine at both positions 105 and 114).
<table>
<thead>
<tr>
<th></th>
<th>Ever-</th>
<th>Mean (SD)</th>
<th>Never-</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GSTM1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1</td>
<td>89</td>
<td>59.5 (12.7)</td>
<td>101</td>
<td>58.7 (10.1)</td>
</tr>
<tr>
<td>*0</td>
<td>84</td>
<td>60.5 (9.3)</td>
<td>120</td>
<td>60.9 (10.5)</td>
</tr>
<tr>
<td><strong>GSTT1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1</td>
<td>135</td>
<td>60.3 (11.4)</td>
<td>174</td>
<td>60.0 (10.1)</td>
</tr>
<tr>
<td>*0</td>
<td>34</td>
<td>59.2 (10.6)</td>
<td>44</td>
<td>59.5 (11.7)</td>
</tr>
<tr>
<td><strong>GSTP1-105</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>74</td>
<td>59.3 (12.2)</td>
<td>95</td>
<td>60.7 (10.6)</td>
</tr>
<tr>
<td>AG</td>
<td>79</td>
<td>60.1 (9.9)</td>
<td>101</td>
<td>59.5 (10.5)</td>
</tr>
<tr>
<td>GG</td>
<td>19</td>
<td>61.7 (12.2)</td>
<td>29</td>
<td>58.7 (9.0)</td>
</tr>
<tr>
<td>AG+GG</td>
<td>98</td>
<td>60.4 (10.4)</td>
<td>130</td>
<td>59.4 (10.1)</td>
</tr>
<tr>
<td><strong>GSTP1-114</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>143</td>
<td>59.8 (11.5)</td>
<td>201</td>
<td>60.2 (10.3)</td>
</tr>
<tr>
<td>CT</td>
<td>26</td>
<td>60.6 (9.7)</td>
<td>23</td>
<td>57.2 (11.0)</td>
</tr>
<tr>
<td>TT</td>
<td>3</td>
<td>62.3 (5.9)</td>
<td>0</td>
<td>n/a</td>
</tr>
<tr>
<td>CT+TT</td>
<td>29</td>
<td>60.8 (9.3)</td>
<td>23</td>
<td>57.2 (11.0)</td>
</tr>
<tr>
<td><strong>GSTZ1-32</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>77</td>
<td>60.9 (12.2)</td>
<td>102</td>
<td>59.4 (10.2)</td>
</tr>
<tr>
<td>AG</td>
<td>76</td>
<td>58.4 (10.6)</td>
<td>93</td>
<td>60.4 (10.5)</td>
</tr>
<tr>
<td>AA</td>
<td>19</td>
<td>62.5 (8.5)</td>
<td>24</td>
<td>60.5 (11.0)</td>
</tr>
<tr>
<td>AG+AA</td>
<td>95</td>
<td>59.2 (10.3)</td>
<td>117</td>
<td>60.4 (10.6)</td>
</tr>
<tr>
<td><strong>GSTZ1-42</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>139</td>
<td>59.8 (11.5)</td>
<td>179</td>
<td>59.6 (10.6)</td>
</tr>
<tr>
<td>AG</td>
<td>26</td>
<td>60.7 (10.1)</td>
<td>36</td>
<td>61.0 (9.9)</td>
</tr>
<tr>
<td>AA</td>
<td>2</td>
<td>62.0 (4.2)</td>
<td>1</td>
<td>68.0</td>
</tr>
<tr>
<td>AG+AA</td>
<td>28</td>
<td>60.8 (9.8)</td>
<td>37</td>
<td>61.2 (9.9)</td>
</tr>
</tbody>
</table>

Table A-4.5 summarises the comparison of mean age-at-onset. There were no differences between mean age-at-onset between different genotypes among PD cases who ever smoked (all p-value >0.23). Similarly, there were no difference in mean age-at-onset between ever smokers and never smokers with the same genotypes (t-test or one-way ANOVA: all p-value > 0.27).
In a multiplex PCR, presence of GSTM1 and GSTT1 genes yields PCR products of 219bp and 459bp bands, respectively. Null genotypes show absence of PCR product. In each reaction, co-amplification of an internal control (beta-hemoglobin gene, product size 268bp) proves the existence of DNA.

Lane M: 50bp DNA ladder
Lane 1: Homozygous GSTM1-null and GSTT1-null
Lanes 2, 3, 5, and 7: Homozygous GSTM1-null
Lane 6: Homozygous GSTT1-null
Lanes 4 and 8: Both GSTM1 and GSTT1 genes are present
Figure A-4.2 PCR-RFLP analysis of GSTP1 genotypes

Lane M: 50bp DNA ladder
Lane 1: Homozygous GSTP1-105Ile
Lane 2: GSTP1-105Ile/105Val
Lane 3: Homozygous GSTP1-105Val
Lane 4: Homozygous GSTP1-114Ala
Lane 5: GSTP1-114Ala/114Val
Lane 6: Homozygous GSTP1-114Val
Figure A-4.3 PCR-RFLP analysis of GSTZ1 genotypes

Lane M: 50bp DNA ladder
Lane 1: Homozygous GSTZ1-32Lys
Lane 2: GSTZ1-32Lys/32Glu
Lane 3: Homozygous GSTZ1-32Glu
Lane 4: Homozygous GSTZ1-42Arg
Lane 5: GSTZ1-42Arg/42Gly
Lane 6: Homozygous GSTZ1-42Gly
CHAPTER 5: NO INTERACTIVE EFFECT BETWEEN THE CYP2E1 GENETIC POLYMORPHISMS AND SMOKING IN PARKINSON’S DISEASE: A CASE-ONLY STUDY

Authors:

Deng YF, Newman B, Dunne MP, Silburn PA, Mellick GD

Contribution of authors:

The candidate (Deng YF) performed all genotyping, data analysis and wrote the entire manuscript. Other authors contributed to the manuscript in terms of providing feedback on the analyses and initial drafts.

This manuscript will be submitted for publication.
ABSTRACT

Smoking appears to reduce the risk of developing Parkinson’s disease (PD). Inter-
individual genetic differences in metabolic pathways may modulate smoking’s effect. The CYP2E1 gene is involved in metabolising chemical compounds in cigarette smoke. Two polymorphisms ($RsaI$ and $PstI$) in the transcription regulatory region of the CYP2E1 gene have been reported to alter the function of this gene. We conducted a case-only study, in an Australian population, to investigate whether the above two polymorphisms are implicated in smoking’s inverse association with PD. Smoking exposure data were collected in face-to-face interviews using a structured questionnaire. We tested the assumption that CYP2E1 genotypes are independent of smoking exposure in the absence of PD by studying 402 healthy, aged individuals. The commonly reported c1 ($RsaI[+]PstI[−]$) and c2 ($RsaI[−]PstI[+]$) allelic haplotypes were detected. Among 400 PD cases (average age at onset=60.0+/−10.7 years), the frequencies of c1/c1 and c1/c2 genotypes were 94.5% and 5.5%, respectively. Statistical analyses did not find any difference in CYP2E1 genotype frequencies between PD cases who ever smoked compared to those who never smoked (odds ratio for interaction (ORi) = 1.00 (95% CI: 0.39-2.51, p=0.99)). Similarly, no CYP2E1 gene-smoking interactions were detected in relation to age at onset of PD. Our data do not support that CYP2E1 $RsaI$ and $PstI$ polymorphisms play a role in the established inverse association between smoking and risk of PD.

Keywords: Cytochrome P450, CYP2E1, genetic polymorphism, interaction, Parkinson’s disease, case-only study
5.1 INTRODUCTION

Parkinson’s disease (PD) is the most common serious movement disorder and the second most common age-related neurodegenerative disorder [1, 2]. World populations, particularly in developed countries like Australia, are ageing. A recent study has reported that both incidence and prevalence rates of PD have increased in the Australian population [3]. Cigarette smoking has been reported to have an inverse association with risk of PD [55, 131, 132]. However, the underlying mechanism remains unclear. Chemical compounds in cigarette smoke, such as nicotine, carbon monoxide and other bioactive molecules, are suggested to play a role [17, 144, 147, 148]. Inter-individual differences in metabolic pathways may alter individuals’ susceptibility to diseases, particularly those triggered by environmental factors [81, 255, 256]. Recent studies have indicated that genetic factors modulate smoking’s protective effect against PD [92, 94, 97, 186, 257].

In humans, chemical compounds are generally metabolised by phase 1 and phase 2 metabolic enzymes. CYP2E1 is one of the most important phase 1 metabolic enzymes involved in metabolising various small-size chemical compounds, including some of those found in cigarette smoke [258]. Two polymorphisms in the CYP2E1 gene, recognized by the restriction enzymes Rsal and PstI in its transcription regulatory region, have been confirmed [259]. These polymorphisms are located in a putative HNF-1 binding site and play an important role in the regulation of CYP2E1 transcription and subsequent protein expression [260]. Based on the presence [+] or absence [−] of Rsal and PstI recognition sequences at these polymorphic sites, four genetic haplotypes are possible: Rsal[+]PstI[−] (referred as c1), Rsal[−]PstI[+] (c2),
RsaI[+]/PstI[+] (c3), and RsaI[−]/PstI[−] (c4). So far, c1 has been reported to be the most common allelic haplotype and is considered the ‘wild-type’. In-vitro experimental studies have indicated that the c2 allele results in greater transcriptional activity than the more common c1 allele [260]. In liver biopsy specimens, CYP2E1 mRNA expression was three-fold higher for subjects with the heterozygous mutant genotype (c1/c2) as compared to subjects with the wild-type genotype (c1/c1) [261]. Moreover, the half-life of acetaminophen, a CYP2E1 substrate, is influenced by CYP2E1 genotype, with c2/c2 homozygotes exhibiting significant reductions over the wild-type.

The CYP2E1 enzyme is expressed in the brain, specifically in the substantia nigra and striatal blood vessels [214, 262]. Previous studies have found that CYP2E1 is associated with free radical production [213, 263]. The prevalence of c2 was reported to be 2-8% in Caucasians and 14-21% in Asian populations [215, 220, 264-266]. So far, c3 and c4 are rarely seen [264, 265]. Several studies have investigated the implications of these polymorphisms in PD susceptibility [158, 220], with inconsistent results. These studies solely analysed the association between CYP2E1 polymorphisms and PD risk without adequately adjusting for the influence of cigarette smoking. This inadequacy could potentially obscure CYP2E1-dependent effects on PD risk if such exist. In addition, none of those previous studies have investigated the potential interactions between CYP2E1 genotypes and smoking in PD. In order to better understand this issue, we conducted a case-only study, in an Australian Caucasian population, to examine whether these two CYP2E1 genetic polymorphisms may modify smoking’s protective effect against the risk of developing PD.
5.2 MATERIALS AND METHODS

Subjects

We recruited 400 white Caucasian PD cases from the Princess Alexandra Hospital Department of Neurology and private neurology clinics in Queensland, Australia. They were comprised of 230 men and 170 women, both in- and out-patients. Probable or definite PD was diagnosed clinically by a movement disorders neurologist, according to accepted criteria: the presence of three of the following cardinal features: resting tremor, rigidity, bradykinesia, postural instability; or two of these features with asymmetry in tremor, rigidity or bradykinesia. All subjects showed good response to L-dopa therapy. Patients whose cardinal features were consistent with another akinetic rigid syndrome, whose cognitive decline was an early feature in the presentation, or who were unable to complete a structured questionnaire were excluded. The age at onset of our PD cases ranged from 26 to 86 years. Sixty-six patients had early-onset PD (age at onset < 50 years). Unaffected, aged subjects (n=402) who were formerly recruited as controls for other PD studies were analysed to examine the assumption that cigarette smoking is independent of CYP2E1 genotypes. These subjects consisted of 156 spouses of patients and 246 other volunteers recruited from patient neighbourhoods and from community organisations. They were able to complete a structured questionnaire either independently or with assistance. All participants gave appropriate informed consent and provided blood specimens for the purposes of DNA extraction. The project was approved by human research ethics committees at each of the participating institutions.
Smoking and other epidemiological data were collected in face-to-face interviews using a structured questionnaire. Subjects were considered smokers if they had ever smoked at least one cigarette per day for a period of a year or more. The smoking data obtained from study subjects included: the age the subject started smoking; cessation history; cigarettes smoked per day during this time period; and current smoking status. These data allow calculation of cumulative exposure to cigarette smoke in pack-years (20 cigarettes = 1 pack; 1 pack/day for a year = 1 pack-year). Our questionnaire also collected data pertaining to family history of PD (first-degree relatives with PD), age, age at onset of PD in cases, history of regular exposure to toxins (herbicides or pesticides, defined by more than once-weekly exposure for a period of six months or more before the onset of PD).

**Genotyping**

Genomic DNA was extracted from peripheral lymphocytes using standard methods (Wizard Genomic DNA Purification Kit, Promega Corporation, Madison, WI, USA). DNA samples were screened for the presence of the CYP2E1 RsaI and PstI polymorphisms by PCR-restriction fragment length polymorphism (RFLP) analysis using forward primer of 5′-CCAGTCGAGTCTACATTGTCA-3′ and reverse primer of 5′-TTCATTCTGTCTTCTAACTGG-3′. The PCR reaction was conducted in a 25ul reaction volume containing 10ng of purified genomic DNA, 10mM Tris-HCl (pH 8.3), 50mM KCl, 1.5mM MgCl₂, 0.2mM of each of deoxynucleotide trisphosphates, 20ng of each of the primers, and 1 unit of Taq DNA polymerase (AmpliTaq Gold, Applied Biosystems by Roche Molecular Systems Inc, Branchburg, NJ, USA). The reaction mixtures underwent the following incubations: 94°C for 3 min, 35 cycles of 94°C for 15 seconds, 61°C for 30 seconds, and 72°C for 45 seconds, followed by a final
extension of 7 min at 72°C. The PCR reaction generated a 412bp fragment containing the RsaI and PstI polymorphic sites in the 5′ regulatory region located upstream of the CYP2E1 transcriptional start site. After restriction enzyme digestion with RsaI or PstI (Promega Corp., Madison, WI) for 6 hours, respectively, the products were electrophoresed on 2.5% agarose gels, stained with ethidium bromide, and examined and photographed over UV light. The presence of RsaI restriction site yielded two fragments of 294 and 118bp, whereas the presence of PstI restriction site yielded two fragments of 351 and 61bp. To prevent contamination or cross-contamination, we prepared the PCR reaction mixture in a DNA/RNA-free room before adding DNA samples. PCR-RFLP analyses were repeated in 12% of the specimens analysed in this study for confirmation of genotyping results. The results were 100% concordant.

**Statistical analysis**

Chi-square tests were performed to check whether the study population was in Hardy-Weinberg equilibrium. Case-only analyses compared genotype frequencies in PD cases between smokers and non-smokers to obtain the odds ratio for interactive effect (ORi). In this case, ORi and 95% confidence intervals (95%CIs) were calculated by unconditional logistic regression, adjusted for age at onset, gender, family history of PD and toxin exposure. ORi>1 indicates more than multiplicative effects between CYP2E1 gene and smoking exposure, ie, gene-smoking interaction increases PD risk; ORi=1 implies solely multiplicative effects, ie, there is no gene-smoking interaction; and ORi<1 suggests less than multiplicative effects, ie, gene-smoking interaction decreases PD risk. Differences in c1 and c2 allele frequencies were assessed using Pearson’s chi-squared test. All p-values presented are two-sided. The Student’s t-test was used to compare the mean of normally distributed continuous variables such as
the age that subjects started smoking and the age at onset of PD, whereas the Mann-Whitney test was used to compare the median of pack-years. The statistical computer software SPSS (ver. 11.0) was used to perform all statistical analyses.

5.3 RESULTS

The characteristics of our PD cases are shown in Table 5.1. The mean age of cases was 66.8 years with the average onset age at 60.0 years. There were more male patients (male:female = 1.35:1). 43.2% of PD cases were ever-smokers, which were much lower than the rate of 56.3% in general Australian population (Australian National Health Survey, 1995, adjusted to the older distribution of our cases). This yields an OR for smoking and PD of 0.6 (95%CI: 0.4-0.8), suggesting that the lifetime smoking status among our PD cases is lower than the general population, a finding consistent with other studies [55, 131].

In both PD cases and unaffected, aged subjects, RsaI and PstI genotype distributions were consistent with the Hardy-Weinberg equilibrium (p>0.57 and p>0.35, respectively). Among our unaffected aged subjects (mean age: 63.9, sd: 11.4 years old), CYP2E1 genotype frequencies among the ever-smoked were similar to those among the never-smoked (p=0.17, adjusted by age and gender). There was also no relationship between CYP2E1 genotypes and the age that an individual started smoking or the number of pack-years smoked (all p-values >0.61). Therefore, these polymorphisms appear to be independent of smoking exposure in unaffected, aged subjects.
The case-only analyses of the distribution of genotypes among the PD groups of never- and ever-smokers are summarised in Table 5.2. The frequencies of CYP2E1 genotypes (c1/c1 and c1/c2) were similar in PD cases irrespective of smoking status (ORi=1.00, 95%CI: 0.39-2.51, p=0.99). There was no difference in haplotype (c1 and c2) frequencies among PD cases who ever smoked, compared to PD cases who had never smoked (Table 5.2). Moreover, CYP2E1 genotypes and smoking status did not influence the onset age of PD (p=0.79) (Table 5.3).

5.4 DISCUSSION

Understanding how host genetic factors interact with environmental factors in PD can be crucial in clarifying the aetiology and pathogenesis of this disease. It may also assist in identifying targets for prevention and treatment strategies for PD. De Palma and colleagues first investigated potential gene-smoking interactions in relation to this disease. They noticed that smoking’s protective effect was lost among patients with a GSTM1 null genotype [186]. Later on, Kelada and colleagues revealed a significant association with heterozygosity at the GSTP1-Ile105Val locus in smokers, with no difference seen between cases and controls who never smoked [97]. A recent study of our PD cases showed that GSTP1-114 genotypes influence the established inverse association between smoking and PD risk [257].

This study specifically focused on examining CYP2E1 gene-smoking interactions in PD based on a case-only design. This design has been considered to be the most powerful method for studying gene-environment interactions, and particularly applicable for relatively rare diseases, such as PD. Our PCR-RFLP analyses found
that in our 400 PD cases, the c1 and c2 allelic haplotypes were present in approximately 97.2% and 2.8% of participants, respectively. We did not detect any individuals with the c3 or c4 allelic haplotype in either PD cases or unaffected, aged subjects (total n=802), suggesting that the CYP2E1 Rsal or PstI polymorphisms may be completely linked in this Australian population. This is consistent with most previous reports in Caucasian populations [215, 267, 268]. The exception is a study conducted by Liu et al, which has identified the “c3” haplotype in a very small number of Caucasian individuals [264].

Analyses of our data did not find any significant difference in the CYP2E1 genotypes among PD cases who ever smoked, compared to PD cases who had never smoked. No CYP2E1 gene-smoking interaction was found in relation to the age at onset of PD either. We did not further divide cigarette smoking into different categories, such as light, moderate and heavy ‘smokers’, since only 22 PD cases were detected with the c1/c2 genotype. Therefore, we were unable to clarify if there is any gene-smoking interaction among heavy smokers.

A limitation of the case-only approach is the requirement that the examined genotypes and exposures are independent of one another in unaffected individuals. Departures from this independence can have a significant impact on the results obtained using this approach [254]. However, in the current study such effects are unlikely to influence our conclusions, given that we demonstrated independence between CYP2E1 genotypes and smoking in a relatively large number of unaffected, aged subjects (n=402) (OR=0.47, 95%CI: 0.16-1.38, p=0.17). This assured the appropriate application of a case-only design in our study. Moreover, given the number of PD
cases (n=400), the case-only approach provided sufficient statistical power to detect modest interactive effects between genotype and smoking status.

To our knowledge, this is the first study specifically examining the interactions between CYP2E1 genotypes and cigarette smoking in the risk of developing PD. The results do not support that CYP2E1 RsaI and PstI polymorphisms play a role in smoking’s protection against the risk of developing PD. However, care must be taken in concluding that the CYP2E1 locus does not interact with cigarette smoking in PD, since our findings cannot exclude the possibility that other functional polymorphisms in this gene may contribute to such interactions. Moreover, the CYP2E1 c2/c2 genotype is extremely rare in Caucasian populations. We did not detect any individual with this genotype in our Australian population. Therefore, our study cannot determine whether c2/c2 genotypes may influence smoking’s protective effect on PD. To clarify this issue, future studies recruiting larger number of PD cases will be needed.

One further concern is the influence of other environmental exposures on the induction of the CYP2E1 enzyme. Since this enzyme is inducible, it is unclear how this might influence the examination of gene-environment interactions. For example, the CYP2E1 enzyme can be highly induced by alcohol consumption, and smoking is strongly associated with alcohol drinking. We did not collect data related to drinking habits. Therefore, we were unable to address this issue in this paper.
In summary, our data suggest: 1. smoking exposure is independent of CYP2E1 genotypes; and 2. there are no multiplicative interactive effects linking smoking and CYP2E1 genotypes with the risk for PD.

ACKNOWLEDGEMENTS

This work was supported in part by Parkinson’s Queensland Inc., the Brain Foundation of Australia and the Geriatric Medical Foundation of Queensland. The authors would like to thank Dr. Diana Battistutta for assistance with data analysis, and Ms. Nadeeka NW Dissanayaka for proofreading the manuscript.
Table 5.1 Demographic characteristics of PD cases

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>230 (57.5%)</td>
</tr>
<tr>
<td>Female</td>
<td>170 (42.5%)</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>66.8 +/- 9.8</td>
</tr>
<tr>
<td>Mean age at onset (years)</td>
<td>60.0 +/- 10.7</td>
</tr>
<tr>
<td>Never-smoked</td>
<td>227 (56.8 %)</td>
</tr>
<tr>
<td>Ever-smoker</td>
<td>173 (43.2%)</td>
</tr>
<tr>
<td>Pack-years smoked</td>
<td>20.9 +/- 21.6</td>
</tr>
</tbody>
</table>

Table 5.2 CYP2E1 genotype and allele frequencies among never- and ever-smokers in PD cases

<table>
<thead>
<tr>
<th>CYP2E1</th>
<th>Ever-</th>
<th>Never-</th>
<th>ORi (95%CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1/c1</td>
<td>164 (94.8%)</td>
<td>214 (94.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c1/c2</td>
<td>9 (5.2%)</td>
<td>13 (5.7%)</td>
<td>1.00 (0.39-2.51)</td>
<td>+ 0.99</td>
</tr>
<tr>
<td>c1</td>
<td>337 (97.4%)</td>
<td>441 (97.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c2</td>
<td>9 (2.6%)</td>
<td>13 (2.9%)</td>
<td>1.10 (0.44-2.84)</td>
<td>0.99</td>
</tr>
</tbody>
</table>

+ORi: odds ratios for interactive effects between genotype and smoking status, adjusted for gender, age at onset, family history of PD and toxin exposure.

Table 5.3 Age at onset of PD among ever- and never-smokers

<table>
<thead>
<tr>
<th>CYP2D6</th>
<th>Age at onset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ever- Mean (SD)</td>
</tr>
<tr>
<td>c1/c1</td>
<td>164 60.4 (10.4)</td>
</tr>
<tr>
<td>c1/c2</td>
<td>9 58.8 (7.6)</td>
</tr>
</tbody>
</table>
Table A-5.1 CYP2E1 genotypes and smoking status in unaffected individuals

<table>
<thead>
<tr>
<th>CYP2E1</th>
<th>Smoking status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ever smoked</td>
</tr>
<tr>
<td>c1/c1</td>
<td>144 (94.7%)</td>
</tr>
<tr>
<td>c1/c2</td>
<td>8 (5.3%)</td>
</tr>
</tbody>
</table>

Table A-5.1 examines whether cigarette smoking is independent of CYP2E1 genotypes in unaffected individuals. Among these subjects, the CYP2E1 genotype frequencies were similar to those reported in other Caucasian populations. There were no significant associations between genotypes and smoking status (OR=1.3; 95%CI: 0.5-3.5; p=0.49).
CHAPTER 6: CYP2D6 POOR METABOLISERS WHO SMOKE CIGARETTES SHOW A LATER ONSET OF PARKINSON’S DISEASE

Authors:

Deng YF, Gartner C, Newman B, Dunne MP, Silburn PA, Mellick GD

Contribution of authors:

The candidate (Deng YF) performed all genotyping, data analysis and wrote the entire manuscript. Other authors contributed to the manuscript in terms of providing feedback on the analyses and initial drafts.

This manuscript will be submitted for publication.

Note to examiners:

Tables containing results from additional analyses mentioned in the published paper, but for which data were not shown, are included at the end of this Chapter.
ABSTRACT

The aetiology of PD remains unclear, although both environmental and genetic factors may play a role. Cigarette smoking is associated with reduced risk of developing PD but the mechanism is unknown. Genetic variability exists in the various metabolic pathways responsible for the metabolism of chemicals, including those found in cigarette smoke. CYP2D6 is an important member of the cytochrome P450 gene family encoding debrisoquine 4-hydroxylase. This enzyme metabolises both endogenous and exogenous compounds. We investigated the interactive effects of CYP2D6 genetic polymorphisms and cigarette smoking on PD in 542 cases and 570 unaffected subjects. The three common CYP2D6 gene polymorphisms (*3, *4 and *5) which result in CYP2D6 poor metaboliser phenotypes (PMs) were genotyped using novel, nested-PCR methods. Smoking exposure data were collected in face-to-face interviews using a structured questionnaire. Our results suggest that among cases without regular pesticide exposures, CYP2D6 PMs who smoked more than 5 pack-years had a later mean age at disease onset (68.6 years) than those with extensive metaboliser phenotypes (EMs) (61.1 years, p=0.02) and non-smokers (60.5 years, p=0.01). Analysis of aged subjects without PD confirmed that neither smoking status nor CYP2D6 PM status was associated with age itself. Our data suggest that one or more components in cigarette smoke may influence age at onset of PD, particularly in individuals who are slow to metabolise them.

Keywords: Cytochrome P450, CYP2D6, genetic polymorphism, interaction, Parkinson’s disease, age at onset
6.1 INTRODUCTION

The aetiology of PD is still unclear. Both environmental and genetic factors may play a role [44, 48]. Recent studies suggest that genetic variation in metabolic enzyme activities may interact with environmental exposures to modify the risk of PD [96, 256, 257].

Cytochrome P450 (CYP) enzymes are key participants in chemical metabolic pathways. As one of the CYP isoenzymes, CYP2D6 encodes debrisoquine 4-hydroxylase, an enzyme metabolising a variety of chemical compounds including pesticides, nicotine and dopamine. The activity of CYP2D6 is genetically determined. Three genetic polymorphisms lead to CYP2D6 alleles with known deficiency in metabolism: (1) G to A transition at the intron 3/exon 4 junction (CYP2D6*4); (2) a single base-pair deletion in exon 5 (CYP2D6*3); and (3) the deletion of the entire gene (CYP2D6*5). Individuals carrying any two of these alleles have almost undetectable enzyme activity. These individuals are defined as poor metabolisers (PMs). CYP2D6 PMs are estimated to represent around 5 to 10% of Caucasians [82, 269].

Pesticide exposures are generally considered to increase PD risk, whereas cigarette smoking has been noted to be inversely associated with this disease. Previous, significant interactions between CYP2D6 PMs and pesticide exposures have been reported. Individuals with CYP2D6 PM status have an exacerbated risk [96, 256]. Since experimental studies suggest that nicotine may play a role in smoking’s inverse association with PD, while CYP2D6 metabolises nicotine, the CYP2D6 genotype may
potentially alter smoking’s protective effect against PD. No previous reports have specifically examined the potential interactions between cigarette smoking and CYP2D6 genotype. Therefore, in the present study, we screened for interactions between CYP2D6 PM status and cigarette smoking that may influence the age at onset (AAO) of PD.

Smoking and pesticide exposures appear to have opposing effects. There is also the possibility that smoking-pesticide and/or CYP2D6 gene-smoking-pesticide interactions may exist. Hence, for the purpose of focusing on examining CYP2D6 gene-smoking interactions, we further analysed our data by removing individuals who claimed regular exposure to pesticides.

6.2 MATERIALS AND METHODS

Subjects
542 eligible PD cases (men 321, women 221) were recruited from the Princess Alexandra Hospital Department of Neurology, private neurology clinics and PD support groups throughout Queensland, Australia, as consecutive patients, during the period of 1999-2004. Diagnostic and exclusion criteria were based on criteria reported previously [257]. We also selected 570 unaffected, aged subjects, who were originally recruited as controls for other PD studies, for examining whether smoking or CYP2D6 genotypes was associated with age [249-251]. To minimise population stratification biases, we only selected subjects of white European ancestry. All participants gave appropriate informed consent and provided blood specimens for the purposes of DNA extraction.
Data were collected in face-to-face interviews using a structured questionnaire. Ever-smoked was defined as at least one cigarette per day for a period of a year or more. We also collected data on the age at which the subject started smoking; cessation history; average number of cigarettes smoked per day during this time period; and current smoking status. We calculated the cumulative exposure to cigarette smoke in pack-years (20 cigarettes = 1 pack; 1 pack/day for a year = 1 pack-year). Other epidemiological data, including family history of PD (first-degree relatives with PD), age (the age when the subjects were recruited), age at onset (AAO, the age when the patients noticed the first symptom of PD) and exposure to pesticides (regular exposure was defined as weekly exposure to herbicides or pesticides for a period of six months or more), were also collected.

**Genotyping**

Whole blood was obtained via venipuncture. Genomic DNA was extracted from peripheral lymphocytes using standard methods (Wizard Genome DNA purification kit, Promega Inc.). We set up a novel nested-PCR method to detect the CYP2D6 genetic polymorphisms *3, *4 and *5 which theoretically may account for more than 95% of PM individuals. In brief, we designed a pair of primers (forward: 5’-ATAGGGTTGGAGTGTTGTTGT-3’, reverse: 5’-GATGCACTGGTGTCACCTTTT-3’) to amplify the entire chromosomal region of CYP2D6, CYP2D7 and CYP2D8. We then used this PCR product as template (5ul of 1:10 dilution) to amplify the region of CYP2D6 using the following primers (forward: 5’-TGCCGCTTTCGCAAACCAC-3’, reverse: 5’-GGCTGGGTCCGAGGTACCC-3’). Amplifications were carried out in a total volume of 20 µl PCR buffer. A mixture
of 8ul the first PCR product added to 8ul of the second PCR product were analysed by 1.5% agarose gel electrophoresis to detect *5 (having band for first PCR but not for second PCR). The second PCR product was used for restriction enzyme digestion with BstNI and DraIII simultaneously. Digestion products were analysed by 2.5% agarose gel electrophoresis to detect *3 and *4. Reliability and validity of the PCR and digestion methods were assessed through re-conducting the genotype assays using a 10% of DNA samples without mutations and all DNA samples with mutations (both heterozygotes and homozygotes). The results were 100% concordant.

**Statistical analysis:** Chi-square tests were performed to check whether the study population was in Hardy-Weinberg equilibrium. Mean AAOs were calculated using General Linear Models, with adjustment for gender and family history of PD, and were compared using one-way ANOVA. SPSS software (release 11.0) was used to perform all statistical analyses.

### 6.3 RESULTS

The characteristics of PD cases are summarised in Table 1. There were 59.2% males and 40.8% females with an average AAO of 60.2 years (SD=10.2 years). Genotypes were obtained for all subjects. All genotype distributions were in Hardy-Weinberg equilibrium. The CYP2D6 PM frequency was 7.0%, similar to those reported in other Caucasian populations (4-10%) [82, 269].

The comparisons of mean AAOs are summarized in Table 6.2. Among all cases (n=542), there were no differences in AAO between different subgroups, although
there was an indication of a dose-response trend among ever-smokers (p=0.57), which was more evident among those who smoked more than 5 pack-years (p=0.32). Table 6.2 also shows that after removal of cases who had a history of regular pesticide exposure, PMs who ever smoked reported experiencing their first symptoms of PD 4.7-years later AAO than those with the extensive metaboliser genotypes (EMs) (65.3 vs 60.6 years), although it did not reach the level of statistical significance (p=0.23). Further analysis showed that among those who ever smoked more than 5 pack-years, the mean AAO of PMs was significantly greater than that for EMs (68.6 vs 61.1 years, p=0.02) (Table 6.2). This was also greater than the mean AAO for all non-smokers regardless of CYP2D6 genotype (68.6 vs 60.5 years, p=0.01). Among CYP2D6 PMs, those who ever smoked more than 5 pack-years had a mean AAO of 9.1 years later than those who never smoked (p=0.04).

6.4 DISCUSSION

A large number of studies have searched for genetic variations that contribute to the risk of PD. Recent findings suggest that gene-environment interactions may hold the key to uncover risk factors for the majority of idiopathic PD [42, 94, 97, 255-257]. The case-only study design has been suggested as being much more powerful and efficient for detecting interactions [103, 104]. However, the successful application of this method relies on independence between the exposure(s) and genotypes under study. Recently, Zareparsi and colleagues [270] suggested that some genetic factors may have a stronger effect on age-at-onset (AAO) of PD than on the risk of developing PD. Identifying and evaluating these genetic variations may be important in understanding the complex origins of PD.
We originally attempted to apply the case-only approach to examine CYP2D6 gene-smoking interactions in relation to the risk of developing PD. However, our preliminary analysis indicated a probable relationship between CYP2D6 genotypes and cigarette smoking in unaffected individuals. Therefore, we were unable to proceed with the case-only approach in this examination. Our tentative analyses using the case-control approach did not detect any interactive effect. However, the findings from case-control analyses need to be considered cautiously since our unaffected subjects may not represent an entirely appropriate control group for our cases.

The present study specifically analysed the interaction of CYP2D6 PM and smoking in relation to AAO of PD. Potential confounding factors, such as gender and family history of PD, were adjusted by using a General Linear Model. Since analysis of our 570 unaffected, aged subjects suggests that pesticide exposure may be associated with CYP2D6 genotype, and analysis of cases suggests that pesticide exposure may be associated with AAO of PD (these are difficult to prove given the sample size of our study) (unpublished), we also considered that pesticide exposure could be a confounding variable in our analysis.

Analysis of all 542 cases did not suggest any interactive effect between CYP2D6 genotypes and smoking. However, previous reports have found a significant interactions between CYP2D6 PMs and pesticide exposure. Individuals with CYP2D6 PM status have an exacerbated risk [96, 256]. Since there is the possibility that a variety of interactions may exist between smoking, pesticide exposure and CYP2D6 genotype in relation to PD, solely adjusting pesticide exposures may not account fully
for the complex confounding effects. Given these facts, we further took steps to simplify our analysis by removing individuals who claimed regular exposure to pesticides. We found that CYP2D6 PMs who ever smoked showed a later average AAO, compared to EMs who ever smoked and all non-smokers regardless of CYP2D6 genotype.

CYP2D6 is expressed in human brain. Given that cigarette smoking is commonly considered as a protective factor against the development of PD, our findings are particularly interesting. One explanation is that the protective effect of compounds in cigarette smoke (perhaps nicotine) is usually removed by the action of CYP2D6. Compared to PMs, EMs may metabolise these protective chemical compounds at a much greater speed, therefore, decreasing their effect. Thus PMs may experience the extended effects of these protective agents.

Payami and colleagues observed that CYP2D6 PMs live longer [82]. This may lead to an unrecognized age effect influencing our results. However, arguing against this is that analysis of 570 aged subjects without PD showed no effect of genotype on age in this group.

It has also been suggested previously that the smoking effect in PD maybe be due to the fact that smokers do not live long enough to acquire PD. This has been strongly refuted by several prospective studies [128]. Moreover, since our analysis focused on CYP2D6 gene-smoking interactions and compared genotypes (PMs vs EMs) amongst smokers, this issue itself does not influence our main findings.
In summary, our study examined the interactive effect between CYP2D6 genotype and cigarette smoking on the AAO of PD. Our data suggest that one or more components in cigarette smoke may associate with AAO of PD, particularly in individuals who are slow to metabolise them.

ACKNOWLEDGEMENTS

This work was supported in part by Parkinson’s Queensland Inc., the Brain Foundation of Australia and the Geriatric Medical Foundation of Queensland. The authors would like to thank Dr. Diana Battistutta for assistance with data analysis.
Table 6.1. Characteristics of the cases

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>321 (59.2%)</td>
</tr>
<tr>
<td>Female</td>
<td>221 (40.8%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>67.0 +/- 9.6</td>
</tr>
<tr>
<td>Range</td>
<td>33-89</td>
</tr>
<tr>
<td>Age at onset (years)</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>60.2 +/- 10.2</td>
</tr>
<tr>
<td>Range</td>
<td>26-86</td>
</tr>
<tr>
<td>Smoking status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>293 (54.1%)</td>
</tr>
<tr>
<td>Ever</td>
<td>249 (45.9%)</td>
</tr>
<tr>
<td>&lt;=5 pack-years</td>
<td>91 (16.7%)</td>
</tr>
<tr>
<td>&gt;5 pack-years</td>
<td>158 (29.2%)</td>
</tr>
</tbody>
</table>
Table 6.2. CYP2D6 genotypes, smoking and age at onset of PD

<table>
<thead>
<tr>
<th>Cases</th>
<th>Genotypes</th>
<th>Ever-smoked</th>
<th>Mean (SD)</th>
<th>&gt;5 pack-years</th>
<th>Mean (SD)</th>
<th>Never-smoked</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (n=542)</td>
<td>EMs</td>
<td>159</td>
<td>59.7 (10.5)</td>
<td>103</td>
<td>60.5 (11.3)</td>
<td>183</td>
<td>60.9 (10.6)</td>
</tr>
<tr>
<td>Carriers</td>
<td>74</td>
<td>61.6 (8.7)</td>
<td>47</td>
<td>61.5 (7.5)</td>
<td>86</td>
<td>58.1 (10.1)</td>
<td></td>
</tr>
<tr>
<td>PMs</td>
<td>16</td>
<td>61.9 (10.5)</td>
<td>8</td>
<td>63.1 (12.4)</td>
<td>22</td>
<td>60.1 (9.5)</td>
<td></td>
</tr>
<tr>
<td>w/o pesticide+ (n=434)</td>
<td>EMs</td>
<td>127</td>
<td>60.6 (10.3)</td>
<td>81</td>
<td>61.1 (11.4)</td>
<td>153</td>
<td>61.8 (10.0)</td>
</tr>
<tr>
<td>Carriers</td>
<td>53</td>
<td>61.4 (9.0)</td>
<td>32</td>
<td>61.2 (7.0)</td>
<td>75</td>
<td>58.1 (10.2)</td>
<td></td>
</tr>
<tr>
<td>PMs</td>
<td>11</td>
<td>65.3 (7.9)#</td>
<td>7</td>
<td>68.6 (5.3)*</td>
<td>15</td>
<td>59.5 (18.8)</td>
<td></td>
</tr>
</tbody>
</table>

+ Cases without regular pesticide exposures; # compared to EMs (60.6 years), p=0.23; * compared to EMs (61.1 years), p=0.02
Table A-6.1 CYP2D6 genotypes and smoking status in unaffected individuals

<table>
<thead>
<tr>
<th>CYP2D6</th>
<th>Smoking status</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ever smoked</td>
<td>Never smoked</td>
<td></td>
</tr>
<tr>
<td>EMs</td>
<td>171 (62.2%)</td>
<td>191 (64.3%)</td>
<td></td>
</tr>
<tr>
<td>Carriers</td>
<td>88 (32.2%)</td>
<td>82 (27.6%)</td>
<td></td>
</tr>
<tr>
<td>PMs</td>
<td>14 (5.1%)</td>
<td>24 (8.1%)</td>
<td></td>
</tr>
</tbody>
</table>

Table A-6.1 presents CYP2D6 genotype frequencies among unaffected subjects (mean age: 67.1 years, sd: 10.6 years). The frequencies of PMs were 5.1% and 8.1% among subjects who ever- and never-smoked, respectively. Although there were no significant associations between genotypes and smoking status (p=0.25), caution must be taken considering our sample size and the relatively low frequency of CYP2D6 PMs.

Table A-6.2 CYP2D6 genotypes and smoking habit in unaffected individuals

<table>
<thead>
<tr>
<th>CYP2D6</th>
<th>Number of subjects+</th>
<th>Age started smoking (years)</th>
<th>Average number of cigarettes smoked per day</th>
<th>Total pack-years smoked</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMs</td>
<td>167</td>
<td>18.3+/-6.7</td>
<td>18.0+/-15.8</td>
<td>25.5+/-25.4</td>
</tr>
<tr>
<td>Carriers</td>
<td>87</td>
<td>19.0+/-5.8</td>
<td>16.4+/-14.4</td>
<td>24.4+/-28.6</td>
</tr>
<tr>
<td>PMs</td>
<td>13</td>
<td>19.6+/-4.0</td>
<td>10.7+/-12.8</td>
<td>10.8+/-18.6</td>
</tr>
</tbody>
</table>

+Six subjects with smoking status did not provide detailed information on smoking.

Table A-6.2 indicates that CYP2D6 PMs are likely to start smoking later, smoke fewer cigarettes per day and fewer pack-years. Although all the differences do not reach the level of significance (all p-value >0.21), these trends, as well as the data in Table A-6.1, suggest a probable subtle linkage between the CYP2D6 gene and smoking behaviour.
Table A-6.3 Pesticide exposures and age at onset

<table>
<thead>
<tr>
<th>Pesticide exposures+</th>
<th>No. of Cases</th>
<th>Mean +/-SD (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No exposure</td>
<td>252</td>
<td>61.3 +/-10.4</td>
</tr>
<tr>
<td>Irregular</td>
<td>175</td>
<td>60.3 +/-10.7</td>
</tr>
<tr>
<td>Regular</td>
<td>106</td>
<td>57.8 +/-10.3</td>
</tr>
</tbody>
</table>

Table A-6.3 shows that patients who have the history of regular pesticide exposures experienced their first symptoms of PD 3.5 years later than those without the history of pesticide exposures, although the difference doesn’t reach the level of significance (p=0.22). This trend suggests that pesticide exposures may potentially influence the age at onset of PD.
Firstly, the entire chromosomal region of CYP2D6, CYP2D7 and CYP2D8 was amplified using the pair of primers: forward: 5’-ATAGGGTGGAGTGGGTGGT-3’; reverse: 5’-GATGCACTGGTCCAACCTTT-3’. Then, the region of CYP2D6 was amplified using this PCR product as template (5ul of 1:10 dilution) and the following primers: forward: 5’-TGCCGCCTTCGCCAACCACT-3’; reverse: 5’-GGCTGGGTCCGAGGTCACCC-3’. A mixture of 8ul the first PCR product added to 8ul of the second PCR product were analysed by 1.5% agarose gel electrophoresis to detect *5 (having band for first PCR but not for second PCR). The second PCR product was used for restriction enzyme digestion with BstNI and DraIII simultaneously. Digestion products were analysed by 2.5% agarose gel electrophoresis to detect *3 and *4, as is shown in this figure.

Lane A: presence of *1 and *3 alleles (*1/*3)
Lane B: Homozygous *5 (*5/*5)
Lane C and D: presence of *1 and *4 alleles (*1/*4)
Lane E and I: Homozygous *1 (*1/*1)
Lane F: presence of *3 and *4 alleles (*3/*4)
Lane G: negative control (no DNA added)
Lane H: Homozygous *4 (*4/*4)
CHAPTER 7: GENERAL DISCUSSION

7.1 INTRODUCTION

Each manuscript has its own separate discussion section that provides a commentary of the findings in relation to the literature. An interpretation of contributing factors, the limitations of the study and the implications in the public health field have also been addressed. This final chapter summarises the general findings in the three manuscripts and takes a macro-level view of how these findings address the research questions and add to the developing literature of how gene-environment interactions influence the risk of PD. Strengths and limitations of this study, and recommendations for further research are also discussed.

7.2 STUDY FINDINGS IN CONTEXT

This thesis investigated whether there are gene-smoking interactions in relation to PD risk between polymorphisms of four GST genes (M1, P1, T1 and Z1), two CYP genes and cigarette smoking. To our knowledge, this is the first genetic epidemiological study of PD using a case-only design to address specifically this question. Our data suggest: 1. smoking exposure is independent of GSTM1, P1, T1, Z1 and CYP2E1 genotypes; 2. smoking may be, to some extent, associated with CYP2D6 genotypes; 3. there are no multiplicative interactive effects linking smoking and GSTM1, T1, Z1 or CYP2E1 genotypes with the risk for PD; 4. there is a multiplicative interactive effect between smoking and GSTP1 haplotype – particularly for genotypes carrying the 114Val allele; and 5. there is a multiplicative interactive effect between smoking and
CYP2D6 PMs – particularly for people who ever smoked cigarettes more than 5 pack-years.

The following three sections substantially discussed issues related to our general findings.

7.2.1 Association between cigarette smoking and specific genotypes

The relationships between cigarette smoking and different GST genotypes (M1, P1, T1, and Z1) were examined in chapter 4. Percentages of ever- and never-smokers, ages at which individuals started smoking and the number of pack-years smoked were compared for different genotypes. Smoking appears to be independent of these genotypes. Using the same subjects, chapter 5 examined the relationship between cigarette smoking and CYP2E1 genotypes. No association was detected. The reviewed literature has not revealed any study examining this issue. Therefore, based on our data, we have concluded that there was no association between smoking and these genotypes.

Chapter 6 examined whether there is an association between smoking and CYP2D6 genotypes. The analyses indicated a probable association between cigarette smoking and CYP2D6 genotypes in 570 unaffected individuals. Although not statistically significant, CYP2D6 PMs were less likely to become smokers, smoked fewer cigarettes per day and fewer pack-years on average. It can be argued that all these results may be solely artifacts. However, we suspect that an association between smoking and CYP2D6 genotypes does exist since almost all smoking information
points in the same direction. Previously, several studies have examined this issue. The first report from Cholerton et al (1996) did not find an association between smoking status and CYP2D6 genotypes [271]. However, their further analysis indicated that some specific smoking behaviours among smokers could be modified by CYP2D6 PMs [272]. Given that the observed association was only subtle, these workers highlighted the fact that a CYP2D6 gene-smoking association could have been missed if stratified analyses were not employed. However, they also suggested that their result could be only a chance finding due to the limitations of their study.

The interpretation of our results should also be considered cautiously since at least four factors could potentially distort an association if it exists:

First, although we have a relatively large sample size with 570 unaffected subjects, there is still limited statistical power to enable the detection of subtle gene-smoking associations. For example, at the genotype frequency of 10% (which is approximately the frequencies of GSTP1-Ala114Val, GSTZ1-Arg42Gly, CYP2E1, or CYP2D6 PMs), our sample size was only powerful enough (i.e., 80%) to detect an association with an OR greater than 1.9, while finding an association with an OR of 1.5 or less would require more than 1700 subjects (see chapter 3).

Second, for examining the gene-smoking associations, the most appropriate population from which to select unaffected subjects is the general population. Our unaffected subjects were selected by convenience. They were originally collected as controls for other PD studies. Although the use of this convenient source of unaffected subjects was considered justifiable (see chapter 3), a potential selection
bias may exist. Specifically, a substantial portion of our unaffected subjects were volunteers, and this group of people may differ from the general population in terms of human behaviours including smoking.

Third, our unaffected subjects were not originally collected for the purpose of examining gene-smoking associations. We were unable to divide these subjects into different smoking categories according to the potential reasons that they became smokers. Therefore, we were unable to comment if there is any gene-smoking association among some subgroups of cigarette smokers.

Fourth, misclassification of genotypes may exist. For GSTM1 and T1, we categorized individuals into deletion or non-deletion, with the latter category actually consisting of both non-deletions and heterozygous one-allele deletions. The PCR method we applied could not differentiate between the constituents of the latter two categories. Since the homozygous non-deletions may have higher enzyme activity than the heterozygous one-allele deletions, this may influence the examination of gene-smoking associations. A related issue exists for our CYP2D6 work. A recent study by Saarikoski et al reported that CYP2D6 ultrarapid metaboliser genotypes could contribute to the probability of being addicted to smoking [273]. Our study did not differentiate ultrarapid metaboliser genotype from EMs due to the time limitation of a PhD project.

It is reasonable to think that all the above factors potentially limit our ability to detect an association between cigarette smoking and the specific genotypes investigated in our study. However, based on our relatively large sample size in analysis, we believe
it is valid to conclude that none of the genotypes investigated were likely to be strongly associated with cigarette smoking. Nonetheless we are unable to exclude a small or modest level of association between CYP2D6 PMs and cigarette smoking.

7.2.2 Specific gene-smoking interactive effects in relation to PD risk

Chapter 4 investigated the interactions between genetic polymorphisms of GSTM1, P1, T1 and Z1 and smoking in PD. This study found no differences in genotype frequencies between the ever- and never-smokers among cases for GSTM1, T1 and Z1. To date, no other studies have examined the associations between these genotypes and smoking.

This study indicated an interaction between the GSTP1 genotypes and smoking that may influence PD risk. Since GSTP1 is expressed in human brain and the polymorphisms lead to functional change, it is possible that altered GSTP1 gene function may influence the development of neurodegenerative diseases including PD. Further analysis suggested GSTP1 variant carriers (114Val carriers) who smoke are at increased risk for PD. This could be of particular interest because smoking is commonly considered to be protective against the development of PD.

Recently, a case-control study by Kelada and colleagues revealed an association for heterozygotes at the GSTP1-Ile105Val locus in smokers with no difference seen between cases and controls who never smoked. They suggested that genetic polymorphism in GSTP1 might interact with smoking and influence the development of PD [97].
One possible explanation for our data is that the protective effect of particular compounds in cigarette smoke requires activation by normal function of the GSTP1 gene. Loss or significant decrease of GSTP1 gene function would lead to a loss of the protective effect. Alternatively, in the absence of normal GSTP1 function, smoking could, in fact, increase risk of PD. Cigarette smoking is known to generate oxidative agents both directly, through metabolism of the components of cigarette smoke, and indirectly, through nicotine-induced activation of dopaminergic activity. GSTP1-114Val carriers may lack the ability to detoxify this increased load of reactive oxygen species. Future experimental studies may be helpful in testing these alternatives.

Our study also found a dose-effect relationship which was most significant in the group of cases with late-onset PD (AAO > 55 years). This age-dependent effect may reflect a more prominent environmental component to late-onset PD compared to early-onset cases. Further studies are required in this area to clarify how GSTP1 gene-smoking interactions influence the risk for PD.

Chapter 5 examined the interactions between genetic polymorphisms of CYP2E1 and smoking in PD. This study found no difference in CYP2E1 genotype frequencies among PD cases who ever smoked compared to PD cases who never smoked. In addition, no CYP2E1 gene-smoking interaction was found in relation to the age at onset of PD. This study is the first report specifically examining CYP2E1-smoking interactions. Given that the polymorphism frequencies are low among PD cases, we could not further divide our subjects into different categories according to the dosage of smoking. Therefore, we were unable to comment as to whether any gene-smoking
interaction exists among heavier smokers. Recently, several other functional polymorphisms in the CYP2E1 gene have been reported. Further studies are required to clarify whether these polymorphisms may contribute to such interactions.

Chapter 6 focused on examining the interactive effects between CYP2D6 genotypes and smoking in relation to age at onset of PD. This study did not apply the case-only approach to examine the differences in genotype frequencies between the ever- and never-smoked PD cases because we suspect that, to some extent, smoking may be associated with CYP2D6 genotypes. Recently, Zareparsi et al found that some genetic factors may have a stronger effect on age at onset than on the risk of developing PD. They suggest that identifying and evaluating these genetic factors may be important in searching for new approaches to delay the onset [270]. A previous study suggested that smoking delays age at onset of PD [274]. Interestingly, our study found that this protective effect was only evident among CYP2D6 PMs who did not have regular exposure to pesticides.

Our study also raised an issue on the relative efficiencies of applying multivariate modelling versus stratified analyses. Previously, two reports including ours suggested significant interactions between CYP2D6 PMs and pesticide exposure in PD (see appendix A). Individuals with CYP2D6 PM status who were exposed to pesticides have an exacerbated risk compared with unexposed subjects and pesticide-exposed CYP2D6 EMs [96, 256]. When analyzing the CYP2D6 gene-smoking interactions in relation to age at onset, we first included all subjects (n=542) in the general linear model and adjusted for pesticide exposure as a confounding factor. This analysis revealed no interactions between CYP2D6 PMs and smoking. Smoking and pesticide
exposures appear to have opposing effects, and there is the possibility that smoking-pesticide and/or CYP2D6 gene-smoking-pesticide interactions may also exist. Solely adjusting pesticide exposure may not account fully for the complex confounding effects. Therefore, we took further steps to simplify our analysis by removing individuals who claimed regular exposure to pesticides. We found that CYP2D6 PMs who ever smoked more than 5 pack-years showed an average 7.5 years later age at onset of PD, compared to EMs who ever smoked, and an average age at onset 8.1 years later, compared to non-smokers (See Table 6.2).

Chapter 6 suggests that one or more components in cigarette smoke may be associated with age at onset of PD. One explanation is that the protective effect of compounds in cigarette smoke (perhaps nicotine) is usually removed by the action of CYP2D6. Compared to PMs, EMs may metabolise these protective chemical compounds at a much greater speed, therefore decreasing their effect. Thus, PMs may experience the extended effects of these protective agents.

The onset of PD has been noted to be a slow and insidious process. Development of noticeable symptoms can be years after the real commencement of PD for some patients. It may be argued that using age at onset of PD as an endpoint (i.e., dependent variable) is inappropriate due to lack of standardised measurement. Indeed, exact timing of PD onset may be difficult. However, a study on test-retest reliability by Reider and Hubble found that PD patients were consistently able to recall their age at which the initial symptom was noted. The kappa statistics was as high as 0.92 [275]. Moreover, the subjects analysed for gene-smoking interactions in a case-only design were all cases. Biases due to this difficulty in timing would be more likely to be
random rather than systematic. Similarly, this is also relevant for timing and measuring exposures before the onset of PD. Therefore, the findings in Chapter 6 are less likely to be greatly influenced by this issue, although caution should be taken in explaining any findings related to age at onset of PD.

For studies examining gene-smoking interactive effects in relation to PD risk, one further concern is the influence of environmental exposures on the induction of metabolic enzyme genes. For example, the CYP2E1 gene has been reported to be highly induced by alcohol consumption while smoking is highly accompanied by drinking. It is still unclear how this will influence the examination of gene-environment interactions. In this thesis we were unable to address this issue.

7.2.3 Issues related to reliability and validity of data

Differences in the quality of measurement of environmental exposures have led to substantial disagreement among published epidemiological studies examining associations between environmental factors and the risk of PD. It is reasonable to believe that unreliability of environmental exposures can be more problematic in investigating gene-environment interactions due to the increased potential of misclassifications. PD cases have two specific characteristics: First, the majority of patients are aged people. It has been argued that data quality may decline with the increase in age. Second, cognitive dysfunction is a recognized feature of PD. These characteristics suggest that retrospective recall may be unreliable.
However, I believe that this issue is less likely to result in substantial influence on the findings in this thesis. As described in Chapter 3, our exposure assessment was based on face-to-face interviews conducted by two experienced interviewers, rather than self-administrated questionnaires. Analyses by another PhD student (Ms. Coral Gartner) showed that, for all types of constructed variables, high test-retest repeatability is obtained for smoking exposure and other epidemiological data relevant to this thesis. For example, when smoking was assessed as ‘ever/never’, the kappa was greater than 0.80; when pack-years were assessed, the intraclass correlation coefficients was greater than 0.97 [276]. Moreover, the high reliability of estimating smoking exposure based on retrospective recall has also been consistently noted in previous studies [275, 277].

In addition, Gartner’s analyses further indicated that there is no difference in repeatability between PD cases and their age-, gender- and residence-matched controls [276]. The results suggest that PD-specific memory deficits do not have a significant impact on exposure data repeatability.

Furthermore, this thesis applied a case-only design. As a retrospective approach, this design may face the uncertainty as to whether an environmental exposure occurred prior to, or after, commencement of the disease. Since the majority of our PD cases who had ever smoked started to smoke before 25 years of age, and given that the age at onset for cases was generally after 50 years, it is unlikely that the initiation of cigarette smoking occurred after commencement of PD.
7.3 STRENGTHS AND LIMITATIONS OF THE STUDY

Although the non-random selection of our samples limits the generalisability of this study, there are several positive aspects to this work. First, this study explored the influence of gene-smoking interactions on PD. In spite of the multifactorial nature of its aetiology, there are limited data specifically exploring the impact of genetic alleles on environmental risk factors in PD studies. Second, this study explored the validation of a case-only approach using data collected from conventional sources. This preliminary case-only study in an Australian population provides a protocol for the methodology and practicality of similar studies internationally.

The general findings, despite the limitations, provide an explanation as to why many previous studies examining the genetic and/or environmental risk factors in PD may have yielded inconsistent results. For example, if the study consists of many subjects who have a history of regular exposure to pesticides (a reported risk factor for PD), inefficient adjustment for this may lead to a finding of ‘no-association’ between smoking and PD risk. Similarly, the data about probable association between CYP2D6 PMs and cigarette smoking provide useful information that may highlight some of the limitations of previous case-control studies examining the association between these genetic factors and PD that failed to consider smoking as a potential interacting factor. As discussed in Chapter 2, these studies showed great inconsistency.

This thesis further highlights the complexity of smoking’s effect on PD. Although, in general, cigarette smoking is inversely associated with the risk of developing PD, it
potentially increases PD risk among some subgroups of people with special genetic backgrounds, such as GSTP1 variant carriers who smoke (see Chapter 4).

This thesis further highlights the importance of stratified analysis in investigating the aetiology of complex disorders. As discussed previously, an interactive effect could be missed if confounding effects were not successfully excluded through adjustment using multivariate analysis. Similarly, inappropriate definition of subgroups may also mask some effects, particularly when the magnitude of the effects is minor or modest. For example, simply defining a group as ‘ever smoked’ does not allow for the huge variation included in this category. Since smoking's effect may not depend on occasional exposure but on consistent exposure, simply comparing this group with the group of ‘never smoked’ has the potential to miss an effect of cigarette smoking. As seen in Chapter 6, the interactive effect between CYP2D6 PMs and smoking was only significant among those who had ever smoked more than 5 pack-years. Unfortunately, because of the difficulty in recruiting large number of subjects, many previous studies were unable to address this issue.

This thesis applied a case-only design to investigate gene-smoking interactions. While this approach has the advantage of eliminating the biases introduced by using convenience samples for control groups and increases the efficiency in detecting interactions, there are nonetheless disadvantages inherent in the case-only design.

First, a case-only design is unable to examine the potential associations between each individual factor (either environmental or genetic) and PD risk. This precludes the collection of definite information as to whether polymorphisms in these GST (M1, P1,
T1 and Z1) or CYP (CYP2D6, CYP2E1) genes play individual roles in PD in the absence of cigarette smoking. Furthermore, the case-only design cannot detect joint additive effects and will detect these as no evidence of interaction [103, 104].

Second, even if a significant interactive effect is found, it still requires careful judgement about whether the association between the genotype and the disease reflects a true causal linkage. Neither case-only nor case-control designs can distinguish between true susceptibility alleles or those that may be markers of the true susceptibility alleles in linkage disequilibrium at a nearby locus.

Third, for maximising generalisability, the case-only design also needs to follow the rules of case selection as for any case-control studies. Ideally, population-based incident cases should be comprehensively or randomly selected. Unfortunately, we, and most researchers in this area, do not have the resources for population-based case ascertainment, principally because of the low prevalence of PD and hence high cost of finding cases.

Finally, and also most importantly, the application of the case-only design is based on the assumption that exposure is independent of genotype. Departures from this independence can have a significant impact on the results obtained using this approach. Practically, there are some genes that might predispose individuals with the genotype to a higher or lower likelihood of the exposure. However, we consistently assessed this independence assumption before proceeding with the case-only approach.
It should be recognised that the magnitude of all the interactions identified in this study was relatively small. Given that we performed no correction of significance values for multiple comparisons, it is also possible that these results are chance findings. Therefore, these findings are indicative rather than definitive. However, the results generate new hypotheses that are worth investigating in future studies. Additionally, this thesis provides a model for undertaking further studies to explore gene-smoking interactions. As better designed studies become available, it will be interesting to see how the evidence builds for interactions contributing to this common neurological movement disorder.

It should also be pointed out that this thesis only examines the target metabolic enzyme genes individually without further considering the potential combined effects of these genes. Since all the genes investigated are involved in metabolising chemical compounds found in cigarette, if two or more genes are needed for altering smoking’s effect, our analyses of genes individually may miss this type of interaction. In fact, Santt and colleagues previously noted a synergistic effect between CYP2D6 PM and GSTM1 null genotypes, reporting that patients with CYP2D6 PM/GSTM1 (-) genotype showed greater increase in risk compared with those with non-CYP2D6 PM/GSTM1 (+) genotype [222]. This issue was not addressed in our thesis due to the limited statistical power of the available sample size. However, for better understanding of PD, we believe that it is also essential to put some effort in investigating the interactive effects between genetic and environmental factors, as well as interactive effects among genes and environmental factors.
7.4 RECOMMENDATIONS FOR FUTURE RESEARCH

Recommendations for future research to address limitations of, or make extensions to, this study are listed below:

- Further studies examining the gene-environmental interactions need to consider the numbers that are required in the sample to be able to explore such interactions. The findings in this thesis were based on a relatively large sample size with 542 PD cases and 570 unaffected subjects. However, the power to detect gene-smoking interactions for genotypes with prevalence less than 10% is limited. Meta-analysis can be an alternative approach in the future if and when a few more studies become available. The success of such an approach will require consistent definitions of exposures and the collection of information about potential confounders.

- Future studies need to pay more attention to measurement of environmental exposures. Great efforts should be put into the accurate estimation of pesticide exposures. Thus, stratification for environmental exposures could be done for all genotypes. This will enable systematic verification or falsification of the hypothesis that most forms of idiopathic PD are caused by a combination of genetic and environmental risk factors. It is therefore vital to develop valid and consistent methodologies for measuring environmental exposures.

- Further studies should recruit a population-based and randomly selected sample which would ensure findings are more representative of the general
Australian population. Our sample may not represent the general population of people with PD, for example, people at lower social-economic status could be under-represented. These people may also be more likely to be smokers and less likely to quit smoking. Although there were no gender or age at onset differences between our participants and the data reported for previous studies, recruiting the sample purposely and conveniently may introduce selection bias. The findings of this survey might not be replicated among other samples that have been recruited in different ways.

- The case-only approach is a powerful and economic screening tool to identify potential interactions. However, to further examine these interactions, other study designs, such as a well-designed case-control study or a cohort study, should follow up. The research questions of this thesis can be more comprehensively and accurately addressed using a longitudinal study design although this is more costly and hence less practical. The longitudinal studies would help to examine the potential gene-smoking interactions as well as enabling conclusions on causality to be drawn. No such longitudinal study has been conducted in Australia or internationally.

- Registration of aging-related disorders including PD is needed in Australia. Such systematic registration helps ensure availability of relevant recent data. Specifically, aging and its related health problems are expected to cause great economic burden in the near future.
• Accumulating evidence suggests that neither genetic nor environmental risk factors for PD can be examined with mutual exclusivity. Further studies need to examine simultaneously the impact of smoking, pesticide exposures as well as other potential risk factors on PD, and assess the influence of family history of PD as well.

• This thesis only considered a limited number of risk factors that may contribute to the development of PD. It should be recognised that the gene-environment associations and interactions are likely to be more complex than this, and may include many additional influences that were not considered.

7.5 CONCLUSIONS

Well-documented evidence has clearly indicated an increasing trend of both prevalence and incidence rates of PD worldwidely. The epidemic demands an urgent investment in prevention of this severe movement disorder. Effective prevention or intervention of PD undoubtedly relies on understanding of its aetiology, which by its nature is complex. Single main effects have been proven unlikely to account for major proportions of the attributable risk.

The results of this research provide some insight into the potential role of interactive effects between the investigated metabolic enzyme genes and cigarette smoking on the aetiology of PD. The study findings provide arguments for why genetic variables in enzymes involved in the metabolism of environmental chemical compounds are valid candidates for conferring differential risk to the development of PD. Data
presented in this thesis highlight the importance of considering the environment context when examining genetic variability in metabolic pathways. Failure to consider crucial gene-environment interactions in such analysis may result in inappropriate dismissal of important risk-altering factors.

The results of this thesis also provide an indication of the power of a case-only design in detecting gene-environment interactions and the complexity of the application of this approach. This project may help further understand this new strategy to investigate the all-important interactions between genes and environment that constitute the determinants of most common complex diseases.

Single gene-environment interactions detected in this thesis are also likely to be poor oversimplification of the reality. However, much understanding of PD will be gained with the concerted efforts of collaborative researchers, examining well-defined study cohorts of adequate sample size, and with standardised and validated instruments for assessing environmental exposures.
APPENDICES

Appendix A: QUT Human Research Ethics Approval, QUT Health and Safety Approval
Appendix B: Questionnaire One (designed by Dr. George Mellick)
Appendix C: Questionnaire Two (designed by Ms. Coral Gartner)
Appendix D: Information Package and Consent Form
Appendix E: DNA Extraction Method (Promega Corporation, USA)
Appendix F: Relevant Publication not Included in the Thesis
Appendix A

QUT human Research Ethics Approval, QUT
Health and Safety Approval
Date: Mon, 02 Sep 2002 12:10:24 +1000
From: Gary Allen <gx.allen@qut.edu.au>
Subject: Minor changes - 2510H
X-Sender: allengj@pop.qut.edu.au
To: c.gartner@qut.edu.au
X-Mailer: QUALCOMM Windows Eudora Version 4.3.2

Dear Coral

I write further to the requests submitted for minor changes to the project, "Environmental Risk Factors for Parkinson's Disease" (QUT Ref No 2510H).

This is to confirm that the Chairperson of University Human Research Ethics Committee has considered these requests under executive powers.

The Chairperson has approved the requested modification to the research team (inclusion of Deng, Dr Yifu as a named member of the research team - to conduct DNA extraction from the samples).

This decision was based upon the assurances provided and is on the strict understanding that there are no other changes / ethical issues not addressed by the approved project.

This decision is also subject to ratification at the 8
October 2002 meeting of UHREC.

However you are authorised to immediately commence the revised project on this basis.

Please do not hesitate to contact me if you have any further queries in relation to this matter.

Regards

Gary Allen
Secretary, UHREC
x2902

Coral Gartner
PhD Student

QUT School of Public Health
Victoria Park Road
Kelvin Grove Qld 4059
AUSTRALIA

Phone: +61 7 3864 5768
Fax: +61 7 3864 3369
Email: c.gartner@qut.edu.au

Ms Coral Gartner
School of Public Health
QUT Kelvin Grove

27 February, 2002

Dear Ms Gartner,

I write further to your application for expedited ethical clearance for your research project, "Environmental Risk Factors for Parkinson’s Disease" (QUT Ref No 25105H). The decision to award your project expedited ethical clearance was reviewed at the 26 February 2002 meeting of the University Human Research Ethics Committee’s (UHREC).

The Committee resolved to ratify the Expedited Ethical Review Panel’s decisions in relation to your project.

If you were awarded conditional clearance by the Expedited Ethical Review Panel the UHREC has ratified both the clearance and the conditions attached to this clearance (ie they must still be addressed before this project may be commenced).

Please do not hesitate to contact me if you have any further queries in relation to this matter.

Yours Sincerely,

Gary Allen
Secretary, University Human Research Ethics Committee
QUT Secretariat
GPO Box 2434
Brisbane Q4001
Telephone: (07) 3864 2902
Facsimile: (07) 3864 1818
Email: gja.allen@qut.edu.au
Cc: Dr Michael Dunne, School of Public Health
Ms Coral Gartner
School of Public Health
QUT Kelvin Grove

April 17, 2002

Dear Ms. Gartner

At its 16 April 2002 meeting, the University Human Research Ethics Committee considered the additional information / revisions you provided in relation to your project "Environmental Risk Factors for Parkinson's Disease" (Ref No QUT 2510H).

The Committee is satisfied that the information provided addresses its concerns, and has confirmed the full ethical clearance status of this project.

Please do not hesitate to contact me if you have any further queries in relation to this matter.

Yours sincerely,

Gary Allen
Secretary, University Human Research Ethics Committee
QUT Secretariat
Telephone: (07) 3864 2902
Facsimile: (07) 3864 1818
Email: gx.allen@qut.edu.au

Cc: Dr Michael Dunne, School of Public Health
Health & Safety Assessment of Projects / Activities

School of Public Health

This form is a planning tool to assist staff and students at QUT in identifying and controlling health and safety hazards they may be exposed to. This form is to be endorsed by the School of Public Health, Health & Safety Committee, and the Head of School, School of Public Health.

PLEASE READ THESE INSTRUCTIONS BEFORE PROCEEDING.

The form consists of several parts.
Everyone involved in a project or activity must complete Form A.
ONLY complete the other parts if told to do so in Form A.

Health & Safety Assessment of Projects

Form A

This Health & Safety Assessment is to be filled in for all projects or activities undertaken. It is to be completed by the staff/student in conjunction with their lecturer/supervisor involved.

Form A will identify forms requiring completion:

Form B – Health & Safety Assessment, Hazard Identification

Form C – Biological Substances

Form D – Radioactive Substances

Form E – Microbiological Materials

Form F – Field Trip Assessment Form

ONLY TO BE COMPLETED IF REQUIRED

Form G – Review

Only for projects where Form B, C, D, and/or F have been completed, a review is to be done annually or whenever significant changes to the workplace, environment, procedures, plant or equipment of the project are proposed. This review section is to be completed on these occasions.

Form H – Project Closure

Only for projects where Forms B, C, D, E and/or F have been completed, this project closure form is to be completed when the project has been completed/ceases.

All staff and students involved in the project or activity as well as the principal investigator or supervisor must sign all forms where indicated and give them to the Workplace Health and Safety Officer for checking and final approval by the School of Public Health, Health and Safety Committee.
FORM A

Project / Activity Information

Project Number 80034

Staff / Student Name YIFU DING

Date 02/11/10

Section 1

Project or Activity Title:

THE POTENTIAL INTERACTION BETWEEN CIGARETTE SMOKING AND SPECIFIC GENETIC POLYMORPHISMS IN THE AETIOLOGY OF PARKINSON’S DISEASE

Brief Description of Project / Activity Work to be carried out:

1. EPIDEMIOLOGICAL STUDIES
   2. QUESTIONNAIRE TO OBTAIN INFORMATION FROM SUBJECTS
   3. GENETIC POLYMORPHISM ANALYSIS
   4. DNA EXTRACTION FROM PERIPHERAL BLOOD CELLS
   5. REAGULAR PCR METHOD TO EXAMINE GENOTYPES OF BST1, CYP2D6, NAT2 AND MAO-B.

Circle the relevant activity(s) pertaining to this evaluation:

Undergraduate Teaching

Postgraduate Research Project: (PhD) Masters Honours

Staff Research

Staff Consulting / Testing

Student Project

Student Field Exercise

Other …………………….Please Specify …………………………………
Section 2

Endorsements:

This project/work has been examined in consultation with the staff members involved and where appropriate, the Health and Safety Representative/Officer. The hazards associated with the work have been identified and the control measures where indicated have been implemented.

Principal Investigator or Supervisor involved:

<table>
<thead>
<tr>
<th>Name</th>
<th>Number</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Staff involved in the project/ activity:

<table>
<thead>
<tr>
<th>Name</th>
<th>Number</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Students involved in the project/ activity:

<table>
<thead>
<tr>
<th>Name</th>
<th>Number</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pass this signed Form A with any other completed forms as required, to the Principal Investigator or Supervisor who must sign the form and pass it on to the School of Public Health, Health & Safety Officer, currently, Lynette Duplock, O-D627, ex 5798, for endorsement.
Part B

1. Project, Work and Equipment Details

<table>
<thead>
<tr>
<th>Project / Activity Number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Investigator or Supervisor Names: Yifei Dong (PhD student); Michael Dinno, Beth Norman, George Mellick (Supervisors)</td>
</tr>
<tr>
<td>Project / Activity Title: The potential interaction between cigarette smoking and specific genotype polymorphisms in the etiology of pulmonary disease</td>
</tr>
<tr>
<td>Funding Source:</td>
</tr>
</tbody>
</table>

Project Description: First part of the project is epidemiological studies to obtain information from subjects using a questionnaire. Second part is genetic polymorphism examination including ASTT, MTHFR, CFHR and NAT2.

| Estimated Commencement Date: 10/01/02 | Estimated Completion Date (or continuing): 30/06/03 |

External collaborators (other Faculties and non-QUT): Dr. George Mellick (Prince Alexandra Hospital)

<table>
<thead>
<tr>
<th>Location(s) of work, eg field site, building, room, campus:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prince Alexandra Hospital</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equipment types involved, if any:</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR amplifier, centrifuge</td>
</tr>
</tbody>
</table>

159
3. Hazard Identification and Control

<table>
<thead>
<tr>
<th>Hazards Identified</th>
<th>Existing Procedures / Equipment</th>
<th>Proposed Procedures / Equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List hazards as listed in section 2 and give details. Refer to list of hazardous substances (NOHSC: 10005 (1994))</td>
<td>Indicate and additional precautions required controlling the hazard identified. Indicate any medical monitoring equipment.</td>
</tr>
<tr>
<td><strong>Eg</strong> HCl, Noise, excess travel, use of cyanide</td>
<td>Include special building, facilities, personal protective equipment, etc.</td>
<td>Eg white gloves, Audiometric monitoring equipment</td>
</tr>
<tr>
<td>ultraviolet</td>
<td>protective glasses</td>
<td>presence of specific additive</td>
</tr>
<tr>
<td>pipetting</td>
<td>proper isolation</td>
<td></td>
</tr>
<tr>
<td>biological</td>
<td></td>
<td>Use &quot;universal&quot; procedures as per Health dept. Treat all samples as infectious</td>
</tr>
</tbody>
</table>

4. First Aid

Please circle

- Are appropriate first aid items available? Yes No
- Have first aid skills been acquired, if necessary? Yes No
- Has consideration been given to the requirement for after-hours first aid treatment? Yes No NA
Part C

Biological Supplement

1. Details of any infectious (or potentially infectious) material or specimens involved.
   - Does this work involve handling material which is infectious (or may be infectious)?
     Yes ☑ No ☐
   - Does this work involve genetically manipulated organisms?
     Yes ☑ No ☐ (Permit No. ___)

   If Yes, attach the GMAC Project Approval Certificate to this Part D.
   - Are imported biological materials used?
     Yes ☑ No ☐
   - Is there a current AGIS Permit for this work?
     Yes ☑ No ☐ (Permit No. ___)
   - Is there a current UREC Approval for this work?
     Yes ☑ No ☐ (Permit No. ___)

List all microorganisms, viruses, cell lines, primary cultures and body fluids used or stored together with the Biosecurity Risk Group they belong to.

<table>
<thead>
<tr>
<th>Microorganism</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood (not expected to contain infectious agents)</td>
<td>☑</td>
</tr>
<tr>
<td>☑ All blood is to be treated as potentially infectious.</td>
<td></td>
</tr>
</tbody>
</table>

2. Medical Monitoring of Staff or Students Involved in this Work.
   (Fig: antibody titre levels, immunisations
   ☑ Hep B immune status to check.
   See pamphlet on Hep B immunisation recommendation which I have attached.)
6. Comments / Endorsements

Students involved in the project or activity are to sign below:

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeva Pena</td>
<td>D.C.</td>
<td>02/01/11</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
<tr>
<td>Name</td>
<td>Signature</td>
<td>Date</td>
</tr>
</tbody>
</table>

Lecturer or Supervisor involved in the project or activity are to sign below:

This Project / Work has been examined in consultation with the Health and Safety Representative and staff members involved. The hazards associated with the work have been identified and the control measures indicated have been implemented.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Dunn</td>
<td>Signature</td>
<td>5/11/11</td>
</tr>
</tbody>
</table>
Memorandum

Faculty of Health - Research Committee

Date: 26 September 2002
To: Mr Yifu Deng, School of Public Health
From: Secretary, Faculty Research Committee
Re: Bio-safety approval - Procedure for DNA Extraction

Yifu,

The Chair of Faculty Bio-safety Committee has reviewed your application for Bio-safety approval for a DNA extraction procedure as part of your PhD Research project "Risk Factors for Parkinson’s Disease). Automatic approval of the procedure is granted on the following grounds:

- This is a standard procedure conducted within the laboratory facilities of the Princess Alexandra Hospital
- The procedure has received appropriate ethical and biosafety approvals through the relevant authorities at the Princess Alexandra Hospital.

For the faculty records could you please obtain a copy of the procedure guidelines and risk management protocols for the laboratory in which you will be working and forward these to me.

I will advise the University Human Research Ethics Committee of approval for your procedure.

Thanks

Pat Smith
Academic & Research Administrator
phone 3864 5778 e-mail p.smith@qut.edu.au

c:\data\files\fric\biosafetyapproval deng1
Appendix B

Questionnaire one
(Designed by Dr. George Mellick)

This questionnaire is not available online. Please consult the hardcopy thesis available from the QUT Library.
Appendix C

Questionnaire two
(Designed by Ms. Coral Gartner)

This questionnaire is not available online. Please consult the hardcopy thesis available from the QUT Library
Appendix D

Information Package and Consent Form

This article is not available online. Please consult the hardcopy thesis available from the QUT Library
Appendix E

DNA Extraction Method
(Promega Corporation, USA)

This article is unavailable online.
Please consult the hardcopy thesis available from the QUT Library
Appendix F

Relevant Publication not Included in the Thesis

This article is not available online. Please consult the hardcopy thesis available from the QUT Library
REFERENCES


38. Mitchell, I.J., et al., *Neural mechanisms mediating 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-induced parkinsonism in the monkey: relative


157. Straub, R.E., et al., *Susceptibility genes for nicotine dependence: a genome scan and followup in an independent sample suggest that regions on*


